Generalized Notation Notation for Active Inference Models

Jakub Smékal[1] [0000−0003−4989−4968] and Daniel Ari Friedman[1] [0000−0001−6232−9096]

Active Inference Institute, USA https://www.activeinference.institute/
blanket@ActiveInference.institute

Abstract. This paper introduces Generalized Notation Notation (GNN), a novel approach to generative model representation that facilitates communication, understanding, and application of Active Inference across various domains. GNN complements the Active Inference Ontology as a flexible and expressive language for education and modeling, by providing a standardized method for describing cognitive models. In this paper we introduce GNN, and provide a Step-by-Step example of what GNN looks like in practice. We then explore "the Triple Play", a pragmatic approach to expressing GNN in linguistic, visual, and executable cognitive models. By situating GNN within the broader context of cognitive modeling and Active Inference, this work aims to bridge and respect the gaps among different modeling settings. The goal of this work is to facilitate interdisciplinary research and application, ultimately promoting the advancement of the field.

Keywords: Active Inference · cognitive models · model representation · hierarchical cognitive models · Bayesian statistics · generative models.

1 Introduction

1.1 Communicating Active Inference: Challenges and opportunities

In recent years, the cognitive sciences have made significant strides towards understanding the complex nature of human cognition and behavior. One promising approach is Active Inference, a unifying theoretical framework that combines perception, action, and learning in a coherent manner [1] [2]. Despite the potential value of models within this framework, the widespread adoption of Active Inference has been hindered by the lack of a standardized method for effectively representing and communicating them. In this paper, we address this challenge by introducing a novel approach to cognitive model representation called Generalized Notation Notation (GNN), which aims to facilitate communication and understanding of Active Inference models across various domains and settings.

1.2 Comprehensively described cognitive models are key

Developing comprehensive, accessible, reproducible, interoperable cognitive models is crucial for the advancement of the Active Inference field. Without a standardized language for describing cognitive models, researchers experience friction when collaborating, sharing, and building upon existing work. In Active Inference research, models are often conveyed through assemblages of natural language, pseudocode, programming languages, analytical formulas, and pictorial representations. In this paper, we present GNN as a flexible and expressive language tailored for expressing Active Inference models, and encompassing various relevant aspects of languages, including ontology, morphology, grammar, and pragmatics. By leveraging GNN as an Active Inferlingua (or Interlingua, Infralingua, or Supralingua), we aim to bridge and respect the gaps among different modeling approaches in order to facilitate interdisciplinary research.

1.3 Goals and structure of the paper

In this paper, first we present the current specification of the GNN language and method (Section 2). To provide an example of GNN in action, we provide the GNN representation of a recent step-by-step Active Inference tutorial [3] [4] (Section 3). We point towards the practical utility of GNN by exploring "the Triple Play" – a pragmatic approach to expressing GNN in three distinct modalities: text-based models, statistical graphical models, and executable cognitive models (Section 4). Lastly, we discuss the implications of our findings, in terms of philosophical and taxonomic perspectives on the diversity of Active Inference generative models, and suggest future research directions (Section 5). By situating GNN within the broader context of hierarchical cognitive models and Active Inference, we hope to inspire further exploration and development of this promising approach to understanding and communicating complex cognitive processes.

2 Active Inference Linguistics: Ontology, Morphology, and Grammar

2.1 Generalized Notation Notation (GNN) overview

GNN describes Active Inference models with ASCII letters and punctuation, structured in a source file that accords with the principles of Markdown. In this section we provide a snapshot of the current specification of GNN.

Updated information on GNN can be found at Github [10] or in Coda [11].

2.2 GNN punctuation

Symbol	Meaning	Example use	Meaning of example
,	A comma is used to separate items in a list.	X,Y	List with X and Y as elements
_	Underscore means subscript.	X_2	X with a subscript 2
=	The equals sign declares the two side of the = to be equal.	X=5	Sets the variable X to value of 5
>	The greater than symbol represents a directed causal edge between two variables.	X>Y	Unidirectional causal edge from X to Y
^	A caret means superscript.	X^Y	X with a superscript Y
-	Hypen is an undirected causal edge between two variables.	X-Y	Undirected causal edge between X and Y
( )	Parentheses are used to bound other expressions.	X^(Y_2)	X with a superscript that is Y with a subscript 2
{ }	Curly brackets are specification of exact values for a variable.	X{1}	X is a scalar with value of 1
[ ]	Rectangular brackets define the dimensionality, or state space, of a variable.	X[2,3]	X is a matrix with dimensions (2,3)
#	A hashtag signals the title header in the Markdown file.	## Title123	Has “Title123” as model title
##	A double hashtag signals a new section in the Markdown file.	## Section123	Has “Section123” as a section name
###	A triple hashtag is a comment line in the Markdown file.	### Comment123	Has “Comment123” as a comment

2.3 GNN source file structure

GNN section	Section meaning	Controlled terms
Image from paper	*Shows the image of the graphical model, if one exists.*	
GNN version and flags	*Describes which specification of LanGauge is being used.*	
Model name	*Gives a name for the model being expressed.*	
Model annotation	*Plaintext caption or annotation for the model. This is a free metadata field which does not need to use any formal or controlled language.*	
State space block	*Describes all variables in the model, and their state space (dimensionality).*	
Connections	*Describes edges among variables in the graphical model.*	
Initial parameterization	*Provides the initial parameter values for variables.*	
Equations	*Describes any equations associated with the model, written in LaTeX. These equations are at least rendered for display, and further can actually specify relationships among variables.*	
Time	*Describes the model's treatment of Time.*	"Static" → Is a static model. "Dynamic" → Is a dynamic model. DiscreteTime=X_t → Specifies X_t as the temporal variable in a discrete time model. ContinuousTime=X_t → Specifies X_t as the temporal variable in a continuous time model. ModelTimeHorizon=X → Specifies X as the time horizon for finite-time modeling.
ActInf Ontology annotation	*Connects the variables to their associated Active Inference Ontology term, for display and model juxtaposition.*	Variables in this section are associated with one or more terms from the Active Inference Ontology. For example "C=Preference" means that the C variable plays the role of the Preference variable.
Footer	*Closes the file and allows read-in from either end.*	
Signature	*Cryptographic signature block (can have information regarding the completeness or provenance of the .S file).*	

3 Step-by-Step GNN

3.1 A Step-by-Step example of GNN applied to Smith et al. 2022

Here we provide several examples of increasing expressiveness of GNN, as defined in the prior section. The examples are directly drawn from "A step-by-step tutorial on active inference and its application to empirical data" by Smith, Friston, and Whyte 2022 [3] [4]. Just as the goal of the initial step-by-step paper was to start simple and progressively add model features till one has arrived at a full Active Inference generative model, here we use that exact same progression to demonstrate the flexibility of GNN.

GNN section	Dynamic Perception	Dynamic Perception with Policy Selection
GNN version and flags	## LanGauge v1	## LanGauge v1
Model name	# Dynamic perception v1	# Dynamic perception with Policy Selection v1
Model annotation	## Model annotations Dynamic Perception This model relates a single hidden state, to a single observable modality. It is a dynamic model because it tracks changes in the hidden state through time.	## Model annotations Dynamic Perception Action Variational Free Energy This model relates a single hidden state, to a single observable modality. It is a dynamic model because it tracks changes in the hidden state through time. There is Action applied via pi.
State space block	## State space block D[2,1,type=float] B[2,1,type=float] s_t[2,1,type=float] A[2,2,type=float] o_t[2,1,type=float] t[1,type=int]	## State space block A[2,2,type=float] D[2,1,type=float] B[2,len(π), 1,type=float] π=[2] C=[2,1] G=len(π) s_t[2,1,type=float] o_t[2,1,type=float] t[1,type=int]
Connections	## Connections among variables D-s_t s_t-A A-o s_t-B B-s_t+1	## Connections among variables D-s_t s_t-A A-o s_t-B B-s_t+1 C>G G>π
Initial parameterization	## Initial Parameterization	## Initial Parameterization
Equations	## Equations s_{tau=1}=softmax((1/2)(ln(D)+ln(B^dagger_tau*s_{tau+1})+ln(trans(A)o_tau) s_{tau>1}=softmax((1/2)(ln(D)+ln(B^dagger_tau*s_{tau+1})+ln(trans(A)o_tau)	## Equations s_{pi, tau=1}=sigma((1/2)(lnD+ln(B^dagger_{pi, tau}s_{pi, tau+1}))+lnA^T*o_tau) s_{pi, tau>1}=sigma((1/2)(ln(B_{pi, tau-1}s_{pi, tau-1})+ln(B^dagger_{pi, tau}s_{pi, tau+1}))+lnA^T*o_tau) G_pi=sum_tau(A*s_{pi, tau}*(ln(A*s_{pi, tau})-lnC_tau)-diag(A^TlnA)*s_{pi, tau}) pi=sigma(-G)
Time	## Time Dynamic s_t=DiscreteTime ModelTimeHorizon=Unbounded	## Time Dynamic s_t=DiscreteTime ModelTimeHorizon=Unbounded
ActInf Ontology annotation	## Active Inference Ontology A=RecognitionMatrix B=TransitionMatrix D=Prior s=HiddenState o=Observation t=Time	## Active Inference Ontology A=RecognitionMatrix B=TransitionMatrix C=Preference D=Prior G=ExpectedFreeEnergy s=HiddenState o=Observation π=PolicyVector t=Time
Footer	# Dynamic perception v1	# Dynamic perception with Policy Selection v1
Signature		

4 Expressing GNN: The Triple Play

4.1 Every GNN expression has a pragmatic and epistemic context

As with communication more broadly, it is key to consider the pragmatic and epistemic context associated with the use of any given Generalized Notation Notation (GNN) expressions. Here, the pragmatic context refers to the practical aspects of communication, such as the preferences and behavioral policies of the communicators, while the epistemic context pertains to the beliefs of those involved in the "ecosystem of shared intelligence" [5]. By taking agent-scale pragmatic and epistemic contexts into account, GNN expressions can be tailored to effectively convey complex concepts using a linguistics that translates seamlessly across different modalities.

Various strategies and tactics may be helpful when employing GNN expressions in different settings and using a spectrum of model precisions (e.g. from informal conversation, to a beautiful presentation, to a fully-documented reproducible research product). Model precision should be balanced against the need for comprehensibility and ease of communication.

Broadly, strategic considerations for expressing GNN include the overall approach to communication, such as the choice of modality and level of technical rigor applied for a given audience. Tactical considerations for expressing GNN include specific techniques for enhancing clarity and understanding, such as using visual aids, examples, and analogies.

Below we highlight the "Triple Play", gesturing towards the high-fidelity rendering of a GNN expression (of a small motif or complete ecosystem-scale generative model) across plain-text, graphical, and executable (computational) forms.

4.2 Text-based models

At its core, GNN is a text-based model that can be rendered into different formats, including mathematical notation, visualized figures, natural language descriptions, algorithmic pseudocode, and executable simulations. The plain-text basis of GNN provides a flexible framework for communicating these computational models, allowing for the integration of different representations to better convey the underlying concepts and relationships of active inference linguistics. Additionally the plain-text basis of GNN enables the use of tools such as Regular Expressions and Large Language Models.

4.3 Graphical models

Graphical models can be visualized using GNN. In the context of Bayesian statistics, graphical models offer a powerful way to represent complex relationships and dependencies. GNN enables the detailing and rendering of clear and informative visual representations that can be easily understood by different audiences. Here GNN can be applied as *post hoc* documentation, by deriving plain-text of

Figures drawn in papers. Additionally GNN can be used *pre hoc* or *in medias res* during the process of designing and implementing Active Inference models. By incorporating graphical models into GNN expressions, we can on one hand utilize the power of Bayesian statistics, and on the other hand benefit from improved visual communication of intricate cognitive models and concepts.

4.4 Executable Cognitive models

Lastly, cognitive models can be executed using GNN. Cognitive models represent the mental processes and structures described by Active Inference. GNN provides a means to specify these models in a formal and precise manner, allowing for their implementation and testing in computational simulations. Critically, GNN as a pseudocode does not restrict which programming language or package ultimately implements the particular generative model in question. As already multiple software implementations of Active Inference exist with only more on the horizon [6], GNN will aid the backwards- and forwards-compatibility of the field. This flexible capability of GNN will enable researchers to explore the implications of various cognitive models, advancing understanding of active inference and catalyzing applications across diverse domains.

5 Discussion

5.1 Summary of key findings and developments

In recent years, the field of cognitive science has experienced significant epistemic advancements and pragmatic developments, especially in the context of Active Inference generative models [2]. These models, which encompass wide ranges of statistical and symbolic model types [7], have been instrumental in shaping our understanding of cognitive processes.

5.2 Philosophical and Taxonomic perspectives on Active Inference model diversity

Ongoing discussions among various factions (e.g. scientific realism and instrumentalism) have raised important questions about the nature, function, and purpose of Active Inference models [8].

The realism-instrumentalism distinction plays a crucial role in shaping our understanding of generative models in Active Inference. Scientific realism posits that these models are approximations of universal truths about reality, implying that they provide veridical representations of the cognitive processes they aim to describe. On the other hand, scientific instrumentalism suggests that these models are intellectual structures that facilitate predictions and problem-solving in specific domains, without necessarily corresponding to any underlying reality. This distinction raises the question of how these models might be viewed as tools for representing objective cognitive truths, and/or for subjective understanding and prediction.

There are significant implications of the realism-instrumentalism conversation in the context of Active Inference cognitive models. Both the realist and instrumentalist perspectives can inform the development of more comprehensive and accurate models, by emphasizing the need for a balance among factors such as explanatory power and predictive utility. By taking into account the strengths and limitations of specific perspectives as applied to particular generative models, researchers can develop methods that are better suited to capture and respect the complexity of cognitive processes, while also providing useful frameworks for problem-solving and prediction.

5.3 Some future research directions

Various future research directions have emerged related to GNN, and will be the subject of ongoing work at the Active Inference Institute and elsewhere. Some salient directions are briefly listed here. In terms of realizing the utility of GNN, improved automatic rendering software and better integration with the Active Inference Ontology [9] will enable the "Triple Play" outlined above. Complex Systems Engineering frameworks such as cadCAD will be useful for specifying the execution order of GNN expressions [6] as well as for mapping the complete landscape of the particular model in question (parameter sweeps across the state-space/belief configurations), for example in the context of cyberphysical systems. From a linguistics perspective it would be interesting to explore possible grammatical case systems, morphology, and dialects associated with GNN. Better integration with natural language processing and formal semiotic methods will enable new kinds of analyses on, with, and for the generative models described by GNN.

5.4 Final thoughts on GNN and Beyond

By adhering to scientific principles such as rigor, accessibility, and plurality, Active Inference researchers can develop and use versatile methods that point towards the intricacies of cognitive processes while also providing practical solutions to real-world problems. As the field continues to evolve, it is crucial to remain open to new ideas, methodologies, action policies, and inferred beliefs, in order to foster a more comprehensive understanding of minds and their workings [1].

References

1. Friston, KJ.: The free-energy principle: a unified brain theory?. Nature Reviews Neuroscience 11, pages 127–138 (2010) https://doi.org/10.1038/nrn2787
2. Parr, T., Pezzulo, G., Friston, KJ.: Active Inference: The Free Energy Principle in Mind, Brain, and Behavior. MIT Press (2022) 9780262045353
3. Smith, R., Friston, KJ, Whyte, CJ: A step-by-step tutorial on active inference and its application to empirical data. Journal of Mathematical Psychology, Volume 107 (2022) https://doi.org/10.1016/j.jmp.2021.102632

4. Smith, R., Whyte, CJ., Murphy, M., Friedman, DA: ActInf ModelStream 001.1, 001.2, 001.3, 001.4: "A Step-by-Step Tutorial on Active Inference". Active Inference Journal (2022) https://doi.org/10.5281/zenodo.7452789
5. Friston, KJ., Ramstead, MJD, Kiefer, AB, Tschantz, A, Buckley CL, et al.: Designing Ecosystems of Intelligence from First Principles. arXiv:2212.01354 (2022) https://doi.org/10.48550/arXiv.2212.01354
6. Friedman, DA, Applegate-Swanson, S, Balbuena, JA, Choudhury, A, Cordes, RJ: An Active Inference Ontology for Decentralized Science: from Situated Sensemaking to the Epistemic Commons. Zenodo (2022) https://doi.org/10.5281/zenodo.7484994
7. Cloutier, J-F: Towards a symbolic implementation of Active Inference for Lego robots. Zenodo (2022) https://doi.org/10.5281/zenodo.6862626
8. Andrews, M: The math is not the territory: navigating the free energy principle. Biology Philosophy 36, 30 (2021) https://doi.org/10.1007/s10539-021-09807-0
9. Active Inference Institute: Active Inference Ontology. (2022) https://doi.org/10.5281/zenodo.7430333
10. Active Inference Institute: Generalized Notation Notation (GNN) Github repo: github.com/ActiveInferenceInstitute/GeneralizedNotationNotation https://doi.org/10.5281/zenodo.7803314
11. Active Inference Institute: Generalized Notation Notation (GNN) Coda: https://coda.io/@active-inference-institute/generalized-notation-notation