Towards Qualitative Spatiotemporal Representations for Episodic Memory

Will Hancock and **Kenneth D. Forbus** and **Tom Hinrichs**[1]

Abstract. Episodic memory is crucial to an intelligent agent's ability to cope with a wide array of diverse tasks. Following Hayes' notion of histories, we hypothesize that qualitative spatiotemporal representations play an important role in organizing episodic memories about continuous phenomena. This paper describes work in progress on constructing qualitative spatiotemporal representations for Freeciv, a strategy game, as a testbed for a more general model of history-based episodic memories.

1 INTRODUCTION

Episodic memory is a persistent contextualized store of specific events (Tulving, 1983), often involving both spatial and temporal aspects. It is a powerful cognitive mechanism, since it supports learning by experience, e.g. using analogy to understand what one's options are when making a decision based on prior experiences. Cognitive systems research has focused on representations that are either completely task independent (Laird & Derbinsky, 2009), (Menager & Choi, 2016), or task specific (Brom et al., 2007). By contrast, Hayes' notion of *histories* (Hayes, 1978), (Hayes, 1989) proposed a general framework for continuous domains. The idea is to use the spatial aspects of the entities involved in an event or situation, combined with their temporal duration, to provide a representation for changes over time. Histories have been heavily used in qualitative reasoning (Forbus, 2019), and thus we hypothesize that they will prove invaluable for organizing episodic memories for continuous aspects of domains.

This paper describes our work in progress on developing qualitative spatiotemporal representations for episodic memories, using the strategy game Freeciv as a testbed. We begin by providing a brief summary of Freeciv and of CogSketch, which we use as a source of spatial representation capabilities. Then we discuss the kinds of histories needed in Freeciv, focusing on the *footprint* of a civilization. We argue that footprints can be used in at least three kinds of strategic problems in playing the game. We close with a summary and discussion of future work.

2 BACKGROUND AND RELATED WORK

We begin by giving an overview of Freeciv and CogSketch, as well as the Freeciv/CogSketch interface.

[1] Computer Science department, Northwestern University, USA, email: wwhancock@u.northwestern.edu

2.1 Freeciv

Freeciv is an open-source strategy game based on Sid Meier's Civilization II. The player is tasked with simultaneously developing a multi-city civilization, managing its growth, economic vitality, and scientific progress, while at the same time developing a military for defense. The game is won when either a player succeeds in getting a colony ship to Alpha Centauri or when a player conquers the world.

Freeciv is an excellent domain for AI research due to its complexity. A typical game board consists of 4,000 tiles, with varying types of terrain. Games typically last for hundreds of turns, with many decisions must be made in each turn. Some are global, across the entire civilization, such as the tax rate, next technology to research, and diplomacy. Workers must be kept productively busy, which can include producing items and building improvements. Military units must be produced as needed, defenses set up and maintained, and when at war, conduct attacks on opponents. By contrast, Go is played on a 19x19 grid, whose spatial properties are uniform, cannot be changed, and are always visible from the start. In addition, each turn in Go only involves a decision to place one piece.

Freeciv is especially useful for exploring episodic memory because important behaviors happen at multiple grain sizes. For example, there is typically an expansion phase, where a player builds out multiple cities, to stake out desirable territory and deny it to competitors. Wars can cause the expansion or contraction of a player's civilization, depending on their success.

2.2 CogSketch

CogSketch (Forbus et al., 2011) is a sketch understanding system that provides a model of high-level visual processing. It provides multiple, hierarchical levels of visual representation, including decomposing digital ink into edges, combining edges into entities, and grouping based on gestalt principles.

The basic level of organization for ink in CogSketch is the *glyph*, one or more ink strokes that are taken to represent some entity (abstract or concrete) in the sketch. Glyphs can be generated by people using a pen or mouse to produce digital ink. But glyphs can also be automatically produced via visual analysis of images (Chen et al. 2019).

CogSketch is capable of computing various spatial relationships between glyphs, including adjacency, relative position, topological relationships and relative size. Properties of individual glyphs can also be computed such as shape attributes (roundness) as well as glyph decompositions such as shape skeletons and Voronoi diagrams. Moreover, CogSketch can decompose glyphs into edges, and group visual entities based on gestalt properties. These capabilities have enabled it to model a variety of visual problem-solving tasks, including Ravens' Progressive Matrices, one of the most common tests used to measure human fluid intelligence. The CogSketch model uses analogical reasoning at multiple levels of visual representations, including re-representing visual descriptions automatically as needed. Its performance places it in the 75th percentile for American adults, better than most adult Americans (Lovett & Forbus, 2017).

Figure 1: A Freeciv map and its CogSketch representation. Pictured are multiple footprint glyphs (red, yellow, etc.) for different civilizations located on a continent glyph with its medial axis decomposition shown.

2.3 The Freeciv/CogSketch Interface

We interface CogSketch with the Freeciv simulator via a Lisp-based API that we constructed. Layers in the sketch are populated with glyphs corresponding to entities from the Freeciv world, such as units and cities. We build upon the spatial representations developed by McLure (McLure & Forbus, 2012) on qualitative representations of regions in the map based on three important properties. First, *continent blobs* consist of connected pieces of land (including any internal lakes). *Terrain blobs* consist of adjacent groups of tiles that share the same terrain category (e.g. grassland, dessert, etc.) *Trafficability blobs* indicate regions where a unit of a particular type can travel. These qualitative regions were shown to be capable of grounding geospatial terms like island and isthmus, learned via analogical generalization (McLure & Forbus, 2012).

2.4 Related Work

Our work on episodic memory to date has focused on learning about primitive actions and their effects, e.g. surprising results of actions (Forbus & Hinrichs, 2018) and durations of actions (Hancock et al., 2018). While several other efforts have explored building agents for Freeciv, none of them use qualitative spatial representations.

There is now psychological evidence that the shape skeletons produced by the medial axis transform are good representations for how people think about 2D shapes (Lowet et al., 2018). This makes it more likely that the qualitative representations CogSketch computes will be understandable by people.

3 QUALITATIVE SPATIAL REPRESENTATIONS

Successfully growing one's civilization in Freeciv involves many spatial considerations. Early on, players vie for territorial control over limited land resources. For example, planning the layout of city sites involves a tradeoff between compactness and sparsity. Compact layouts cause resource competition between cities, while sparse layouts increase vulnerability to attacks by hostile nations.

Our goal is to develop region-based qualitative spatial representations to support strategic reasoning and learning in Freeciv. See Figure 1 for examples. For example, comparing results across games should be facilitated by history-based episodic memories. If a player tries a new tactic for more rapid growth, success could be measured by comparing either how far each civilization got by a particular turn, or by how many turns it took to reach a desired size.

3.1 Encoding Principles

We seek to construct region representations that are concise, sparse, and local. *Concise* means that the representations respect the relevance principle, that is, they do not make unnecessary distinctions. Regions where a relevant property is constant should be one entity, and subdivided only when required because of some other important constraint. *Sparsity* is an aid to learning. There is always a tradeoff in the amount of information encoded: too little, and there isn't enough signal to learn the appropriate distinctions. Too much, and learning is made more difficult because the space of hypotheses is larger. We use the Structure-Mapping Engine (SME; (Forbus et al., 2017)) to do comparisons and in analogical retrieval and generalization. More entities means more potential relations between them, which means more work to match descriptions. All else being equal, keeping the number of entities small is preferable. This is one reason that qualitative representations are better than, for example, using tiles or other highly granular quantitative representations as commonly done in reinforcement learning research. *Locality* means that networks of relationships should be computed between entities that are spatially local to one another. This makes the set of relationships computed for two descriptions more likely to match when they are similar. This locality heuristic is already built into several of CogSketch's algorithms, e.g. positional relationships are only computed between glyphs that are adjacent.

3.2 Perceptual Encoding

CogSketch can operate on glyphs to construct new entities dynamically. It can create glyphs from the intersection, difference, or union of two existing glyphs. It can also create new glyphs by intersecting visible glyphs with the drawing pane. CogSketch can decompose a glyph into a network of junctions and edges. Just as there are many attributes and relationships that CogSketch can compute at the glyph level, there is also a collection of qualitative spatial relationships and attributes that can be automatically computed at the edge level. These include connectivity, relative length, concavity, axis alignment, and curvature. See (Lovett & Forbus, 2011) for a full catalog of edge-level descriptors.

Following (McLure & Forbus, 2012) (Lowet et al., 2018), we use the medial axis transform on blob boundaries to compute shape skeletons, since these are psychologically plausible predictors of shape similarity, and hence should aid analogical retrieval. The medial axis transform is a directed connected graph. Edges are directed by the sign of the derivative of the blob radius. Source junctions occur at vertices when all edges are outgoing. Sink junctions occur at vertices when all edges are incoming. Moreover, the shape skeleton supports multiple kinds of reasoning, as will be outlined in Section 3.4.

3.3 Footprints

In addition to the continent, terrain, and trafficability blobs of (McLure & Forbus, 2012), we introduce a new blob, the *footprint* of a compound entity. In Freeciv, a civilization is a compound entity. Outside the game, examples of compound entities include groups of people, groups of vehicles, and ant colonies. The footprint is the portion of terrain that includes the entities of the compound entity and the interior spaces within it. In Freeciv, a city has a footprint of the 16 tiles around it. The civilization's footprint can be considered the convex hull of the city footprints (see Figure 2). Since a civilization can have cities on more than one continent, the footprint of that civilization can span continents (and portions of ocean between them) as well. Depending on the particulars of city placement, there can even be partial overlap (i.e. RCC8's PO relation (Cohn et al., 1997)) between footprints of two civilizations. The footprint is represented as a glyph in CogSketch. Cities whose footprints lie on the boundary (and hence RCC8 TPP, i.e. tangential proper part) are *border* cities, and all other cities are *internal* cities.

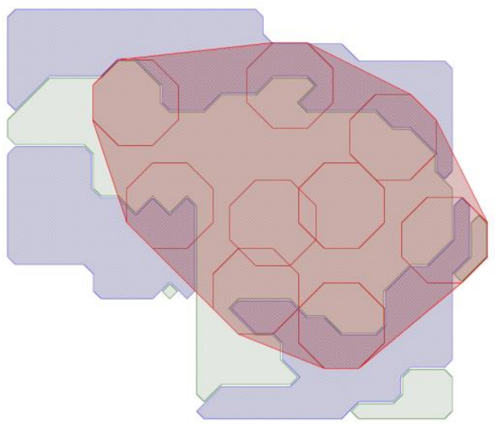

Figure 2: A footprint region consisting of the convex hull of a set of city footprints

3.4 Reasoning with Footprints

We focus our attention on three tasks within the game: civilization expansion, enemy incursion monitoring, and city specialization.

Civilization expansion is critical early in the game because cities are the wellspring of units and resources. There is a tradeoff between increasing the number of cities versus growing existing cities. Planning the expansion of a civilization requires taking a number of factors into account. First, terrain itself contains resources to be mined or farmed, and terrain is required to build new cities. Thus, one component of the game is a race to acquire territory. Terrain with ample resources will enable cities to grow more quickly, whereas poor terrain leads to cities that languish. Proximity to potentially hostile nations is another important issue. Expansion, like most aspects of the game, requires assessing multiple factors and finding good solutions to tradeoffs.

We hypothesize that city sites located on the border of a footprint region are more likely to be subjected to enemy invasion. Furthermore, a continent blob medial axis decomposition can lend insight into the value of potential city sites. Border cities that are connected to enemy territory by the shape skeleton should be particularly susceptible. It follows that these cities should choose terrain that facilitates defense, such as on hill and forest terrain, or on tiles with river specials. Additionally, some terrain features (i.e. an isthmus) make ideal defensive city sites. The shape skeleton can be queried for sink junctions between one's footprint region and enemy locations for suitable sites.

The second problem, incursion monitoring, involves detecting potential attackers early enough to do something about them (e.g. move defensive units into position, or build new defenders or other military improvements in a city under threat). A simple qualitative model would be to establish a radius around every city, such that if a unit enters that radius, it gets the system's attention to assess its intent. Our hypothesis is that the footprint, expanded by a trigger radius, will provide a more sensitive incursion detection measure, since the terrain inside the footprint will typically have workers making improvements, settlers heading for new territories, and interior cities that are less heavily defended. Long border stretches between cities are natural candidates to place sentries, to provide early warning of incursions. Thus, a player should benefit from boundaries that are based on territory rather than just on cities.

The third task, city specialization, involves optimizing a civilization's performance by assigning functional roles to cities. Border cities are natural defenders since they are closer to where attackers might come from. Internal cities, since they have to spend less on defense, can be specialized to economic activities (e.g. marketplaces, banks, factories) or scientific research (e.g. libraries, and universities). Cities that lie nearer to interior arterial regions of the shape skeleton can balance between defense and resource production.

4 Experiments

Fifteen Freeciv scenarios have been generated for testing. A Freeciv scenario consists of a randomly generated map along with settings that control map features and enemy nations. For our scenarios, we have kept most of the settings at their default values. We have decreased the default map size from 4000 to 2000 tiles to encourage acts of war. We have disabled huts to normalize start-of-game civilization development. When entered by a unit, huts randomly select an event which can drastically affect the strength of a civilization. They can grant new technologies, new cities, or unleash enemy barbarian hordes. Their removal ensures fairness in gameplay. We also disable fog of war. Fog of war limits visibility to previously explored tiles that no longer have friendly units present. While ultimately we would like our to be able to reason about uncertainty, it is not the explicit goal of these experiments. We are testing the aforementioned three tasks:

4.1 Incursion Detection

Threat detection provides an important decision-making signal to an agent. When facing an impending attack, resources can be moved from growth-based production to defense. We define baseline and spatial threat-detection mechanisms. Our baseline detector signals a threat when enemy combat units are found within a civilization's national borders. National borders are calculated by the underlying game engine. Their magnitude roughly corresponds to the populations of their constituent cities. They start off small and are possibly disjoint if their constituent cities are far enough apart. In contrast, our footprint region consists of the convex hull of all civilization cities. Threat detection for both national borders and footprint regions is signaled when an enemy unit is within their borders.

Additionally, our qualitative incursion detector analyzes enemy unit movement. It flags enemy combat units that are moving towards the footprint region as threats. Enemies within a threshold distance whose distance from the footprint decreases or stays the same over three turns are flagged.

4.2 Civilization Expansion

The default role of a city is to produce settlers, to allow more cities to be founded. Our baseline agent continues to produce settlers in all cities until it detects an enemy incursion. It then switches to producing combat units until all cities are lost, or there is no longer a threat.

Footprint representations allow further reasoning about expansion. Encoded footprint RCC8 relations can assist in reasoning about the availability of city building sites. Blobs that are adjacent to a civilization's footprint that are not owned by another player signal the potential for a civilization to expand. When none of these blobs exist, cities are told to stop producing settlers and move to satisfy other goals (e.g. defending themselves).

4.3 City Specialization

For our baseline condition, all cities produce the same resources. Settlers are built when an enemy incursion is not detected, and combat units are produced otherwise.

For our qualitative trials, cities are specialized based on their spatial properties. Footprint representations are decomposed for their constituent cities. A city is either on the hull or interior of the footprint. When no threat is detected, hull cities adjacent to enemies first ensure their own defense by building units or defensive improvement (e.g. city walls). All other cities build settlers to expand the civilization.

When a threat is detected, all cities ensure their own defenses. When interior cities are sufficiently defended, they are specialized to other tasks (producing gold, science, etc.).

4.4 Trials

Each scenario is run for 100 turns for the baseline and qualitative conditions. Civilization attributes are saved over each turn. We focus on total city count as a metric of success, but also consider treasury reserves and average city size across the civilization.

4.5 Results

Experiments are currently under way.

5 Temporal Components of Episodic Memory Histories

So far we have focused mainly about space, what about time? For these experiments, reasoning is conditioned on two different phases: pre and post-incursion. We hypothesize that flexible agents must be able to reason about many more types of temporal phases.

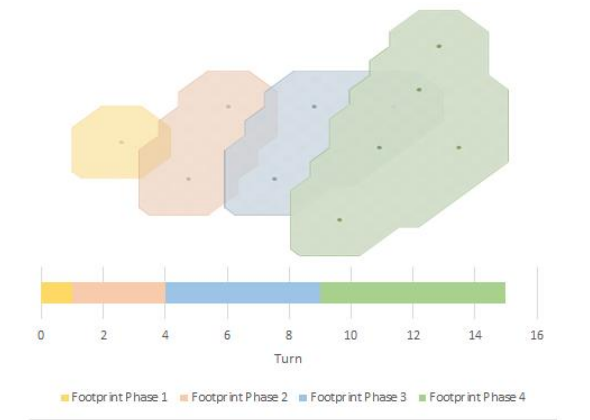

Figure 3: Temporal phases over the expansion of a civilization's footprint

For civilization footprints, a natural representation is to divide time based on when the number of cities in a civilization changes. Since cities are not created (or destroyed) on every turn, this quantizes time into reasonably meaningful chunks. Note that wartime setbacks or economic disasters can reduce number of cities, and it is possible to have both creation and destruction happen in the same turn, so a better measure would be the set of existing cities as a means to individuate temporal intervals. Thus every time this set changed, the previous spatial representation of the footprint and the interval over which it held would be stored as an item in episodic memory, a case that can be subsequently reasoned about. Changes in diplomatic state, e.g. wars, also should trigger the creation of episodes, and in those cases their spatial component might be the footprint of the combatant's military units. Figure 3 provides an illustration, where 15 turns worth of time is quantized into four, based on the cities that exist during those turns. We plan to use several criteria to evaluate episodic memory histories. The first is whether or not it can be used to provide a concise summary of activities within a game. The second is whether or not it can be used to support strategic reasoning, as outlined above. The third is whether or not it can be used to learn improved models of the game's dynamics, strategies and tactics.

6 Discussion

We have described how we are exploring qualitative spatiotemporal representations to make episodic memories in the form of Hayes-style histories. We already have automatic construction of some of these qualitative spatial representations, namely civilization footprints, which we think are promising for supporting strategic reasoning in Freeciv.

Our next step is to extend our Companion cognitive architecture with history constructors, tied initially to Freeciv. We plan to experiment with them, tuning the representations to support summarization, reasoning, and learning. Assuming they prove effective, we will also explore these ideas in other domains, such as story understanding and commonsense reasoning.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Air Force Office of Scientific Research under Award No. FA9550-16-1-0138.

REFERENCES

Brom, C., Pešková, K., & Lukavský, J. (2007). What Does Your Actor Remember? Towards Characters with a Full Episodic Memory. In M. Cavazza & S. Donikian (Eds.), Virtual Storytelling. Using Virtual Reality Technologies for Storytelling (pp. 89–101). Springer. https://doi.org/10.1007/978-3-540-77039-8_8

Chen, K., Rabkina, I., McLure, M., & Forbus, K. (2019) Human-like Sketch Object Recognition via Analogical Learning. AAAI 2019.

Cohn, A. G., Bennett, B., Gooday, J., & Gotts, N. M. (1997). Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. GeoInformatica, 1(3), 275–316.

Forbus, K. D. (2019). Qualitative Representations: How People Reason and Learn about the Continuous World. MIT Press.

Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2017). Extending SME to Handle Large-Scale Cognitive Modeling. Cognitive Science, 41(5), 1152–1201.

Forbus, K. D., & Hinrichs, T. R. (2018). Qualitative Reasoning for Decision-Making: A Preliminary Report. Proceedings of the QR'18, 9–15.

Forbus, K., Usher, J., Lovett, A., Lockwood, K., & Wetzel, J. (2011). CogSketch: Sketch Understanding for Cognitive Science Research and for Education: Topics in Cognitive Science. Topics in Cognitive Science, 3(4), 648–666.

Hancock, W., Forbus, K. D., & Hinrichs, T. R. (2018). Optimizing Strategy Game Planning Via Combining Analogical Learning with Function Learning. Advances in Cognitive Systems, 6, 97–109.

Hayes, P. J. (1978). The Naive Physics Manifesto. University of Essex.

Hayes, P. J. (1989). Naive physics I: Ontology for liquids. In Readings in qualitative reasoning about physical systems (pp. 484–502). Morgan Kaufmann Publishers Inc.

Laird, J. E., & Derbinsky, N. (2009). A Year of Episodic Memory. Workshop on Grand Challenges for Reasoning from Experiences.

Lovett, A., & Forbus, K. (2011). Organizing and Representing Space for Visual Problem-Solving. 8.

Lovett, A., & Forbus, K. (2017). Modeling visual problem solving as analogical reasoning. Psychological Review, 124(1), 60–90.

Lowet, A. S., Firestone, C., & Scholl, B. J. (2018). Seeing structure: Shape skeletons modulate perceived similarity. Attention, Perception, & Psychophysics, 80(5), 1278–1289.

McLure, M. D., & Forbus, K. D. (2012). Encoding Strategies for Learning Geographical Concepts via Analogy. 9.

Menager, D. H., & Choi, D. (2016). A Robust Implementation of Episodic Memory for a Cognitive Architecture. 6.

Tulving, E. (1983). Elements of Episodic Memory. Clarendon Press.