Between min cost search and least action

Tim Fernando

Abstract submitted to a workshop in Edinburgh, April 2025

A recent assessment of large language models claims LLMs have, through pre-training by next-word prediction, largely achieved "formal linguistic competence (knowledge of linguistic rules and patterns)" but not "functional linguistic competence (understanding and using language in the world)" [M$^+$24]. Functional competence, it is suggested, calls for fine-tuning through, for example, reinforcement learning. Fine-tuning is about alignment, which the present paper approaches by refining notions of state around which to carry out a minimum cost search, supported by finite-state methods (and tools such as [Mona]) for a discretized least action principle [F63,MW01].

To see what a notion of state adds to next-word (or token) prediction, recall that the probability of a string $a_1a_2\cdots a_n$ of n tokens a_i is the product of the conditional probabilities $P(a_{i+1}|a_1a_2\cdots a_i)$ of (next token) a_{i+1} given $a_1a_2\cdots a_i$, for $0 \leq i < n$

$$P(a_1a_2\cdots a_n) = \prod_{i=0}^{n-1} P(a_{i+1}|a_1a_2\cdots a_i) \tag{1}$$

(e.g., [RN20]). While it is customary with automata to extract $a_1a_2\cdots a_n$ from a chain

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \cdots \xrightarrow{a_n} q_n \tag{2}$$

of state transitions linking an initial state q_0 to a final state q_n, no state q_i is mentioned in (1) — unless q_i is say, its history $a_1a_2\cdots a_i$ in which case $P(a_{i+1}|q_i)$ appears in (1) as the factor $P(a_{i+1}|a_1a_2\cdots a_i)$. Already with n-grams, other notions of state are at play, suggesting $P(a_{i+1}|q_i)$ as an alternative to $P(a_{i+1}|a_1a_2\cdots a_i)$ for transitions (2) subject to the Markov property; for example, the trigram model

$$P(a_{i+2}|a_ia_{i+1}) = P(a_{i+2}|a_1a_2\cdots a_ia_{i+1})$$

arises when a state, q_{i+1}, can be identified with the two most recent tokens seen, a_ia_{i+1}. Similar reductions for actions a_i in a path (2) from a *Markov decision process* apply to probabilities describing policies

$$P(a_i|q_{i-1}) = P(a_i|q_0a_1q_1\cdots a_{i-1}q_{i-1}) \qquad \text{(how often to do } a_i \text{ at state } q_{i-1}) \tag{3}$$

and transitions between states

$$P(q_i|q_{i-1}a_i) = P(q_i|q_0a_1\cdots q_{i-1}a_i) \qquad \text{(how probable } q_i \text{ is after } a_i \text{ at } q_{i-1}). \tag{4}$$

The product of the policy and transition probabilities is the conditional probability $P(a_iq_i|q_{i-1})$ of doing a_i with the consequence q_i given q_{i-1}. Over $i = 1$ to n, their product is the probability of (2) given q_0

$$P(a_1q_1\cdots a_nq_n|q_0) \ = \ \prod_{i=1}^{n} P(a_iq_i|q_{i-1}) \tag{5}$$

assuming (3) and (4). Mapping probabilities to *surprisals* [S48] as costs

$$c_i \;:=\; -\log P(a_i q_i | q_{i-1}) \qquad \text{for } 0 < i \leq n$$

turns the product in (5) to a sum

$$-\log P(a_1 q_1 \cdots a_n q_n | q_0) \;=\; \sum_{i=1}^{n} c_i \tag{6}$$

of costs c_i. As $P(a_i q_i | q_{i-1}) = P(q_{i-1} a_i q_i)/P(q_{i-1})$, each cost c_i is a difference

$$c_i = K_i - V_i \;\text{ where }\; K_i := -\log P(q_{i-1} a_i q_i) \;\text{ and }\; V_i := -\log P(q_{i-1})$$

between terms K_i and V_i related by marginalization $P(q_{i-1}) = \sum_{a,q} P(q_{i-1} a q)$ and the intuition: given the benefit V_i, pay the cost K_i. To view K_i and V_i as analogous to kinetic and potential energies, and their difference c_i as a cost-minus-benefit Lagrangian $L(q, \dot{q})$ over a state vector q and velocity vector $\dot{q}$, a discretization of mechanics is in order, "the starting point" of which, according to [MW01], is to "regard two nearby points as being the discrete analogue of a velocity vector" (page 360). The main idea of the present work is to reduce the time-step h between "nearby points" by refining the notion of state, whilst recognizing the label on a state transition $q_{i-1} \xrightarrow{a_i} q_i$ in a *generalized velocity* (q_{i-1}, a_i, q_i). The label a_i brings in an agent/policy (3), shaped by rewards and an exploration-exploitation dilemma that go into alignment.[1]

From action as a label on state transitions, we turn to action as an integral over time, $\int_{t_1}^{t_2} L(q, \dot{q}) dt$, discretized as a finite sum. We map a path (2) to the sum $\sum_{i=}^{n} c_i$ in (6), setting aside for the moment the time-step h, which "to relate discrete and continuous mechanics it is necessary to introduce" [MW01, page 370]. (For the sum to align to the integral, let h approach 0.) The interest in (2) reflects the view that "the analysis of the patterns generated by the world in any modality, with all their naturally occurring complexity and ambiguity" is served by "reconstructing the processes, objects and events that produced them" [M94, page 187]. That reconstruction is complicated by the various bounded granularities at which paths (2) are observed (as patterns), leading to strings both more and less detailed than $a_1 a_2 \cdots a_n$. These granularities and strings can be encoded as *signatures* Σ and *Σ-models* of an *institution* [GB92], exhibiting deformations studied in [GM07,M94]. This is outlined in [F23], where time is compressed to ensure change discernible through *generalized coordinates* of a state q (relativized to Σ). These coordinates can be identified with *inner cells* in Kleene's analysis of nerve nets [K56], taking values approximated to finite precision. Refinements within and between institutions effectively reduce the time-step h of the discrete Lagrangian. Now, just as there are multiple Lagrangians L (not all of the form $K_i - V_i$), so too are there many distributions P (from, for example, different policies). Where a signature Σ fails to support assumptions (3) and (4), we redefine c_i as $-\log P(a_i q_i | q_0 a_1 \cdots a_{i-1} q_{i-1})$. Whether or not entropy maximization in say, [S$^+$20] is applicable, we can let the relevant conditional probabilities be uniform (for priors, subject to update from learning), working with a sample space of Σ-models under finite-state constraints to fine-tune Σ and notions that Σ associates with paths (2). (Details in the presentation/paper.)

[1] Observe that (1) mentions *no* goal or final state q_n to search, even if a state q_i is its history. By contrast, the costing (6) of the path (2) from q_0 to q_n informs a "search for links" [W90].

References

F23. T. Fernando. 2023. Triadic temporal representations and deformations. In *Proc. Natural Logic meets Machine Learning IV*, pages 51–61. ACL Anthology.

F63. R. Feynman. 1963. The principle of least action. In *The Feynman Lectures on Physics*, volume 2, chapter 19. Caltech, `https://www.feynmanlectures.caltech.edu/II_19.html`.

GB92. J.A. Goguen and R.M. Burstall. 1992. Institutions: Abstract model theory for specification and programming. *Journal of the ACM*, 39(1):95–146.

GM07. U. Grenander and M. Miller. 2007. *Pattern Theory: From Representation to Inference*. Oxford University Press.

K56. S.C. Kleene. 1956. Representation of events in nerve nets and finite automata. In C.E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 3–42. Princeton University Press.

M^+24. K. Mahowald, A.A. Ivanova, I.A. Blank, N. Kanwisher, J.B. Tenenbaum and E. Fedorenko. 2024. Dissociating language and thought in large language models. *Trends in Cognitive Sciences*, 28 (6).

MW01. J.E. Marsden and M. West. 2001. Discrete Mechanics and Variational Integrators. *Acta Numerica*, 10:357–515.

Mona. The MONA project. `https://www.brics.dk/mona/`.

M94. D. Mumford. 1994. Pattern Theory: a unifying perspective. *First European Congress of Mathematics*, pages 187–224. Birkhäuser.

RN20. S. Russell and P. Norvig. 2020. *Artificial Intelligence: A Modern Approach*, 4th edition, Pearson.

S^+20. Y. Savas, M. Ornik, M. Cubuktepe, M.O. Karabag and U. Topcu. 2020. Entropy Maximization for Markov Decision Processes Under Temporal Logic Constraints. *IEEE Transactions on Automatic Control*, 65(4):1552–1567.

S48. C.E. Shannon. 1948. A Mathematical Theory of Communication. *Bell Systems Technical Journal*, 27(3):379–423.

W90. J.A. Wheeler. 1990. Information, physics, quantum: The search for links. In W. Zurek, editor, *Complexity, Entropy and the Physics of Information*, pages 3–28. Addison-Wesley.