LING5702: Lecture Notes 11
Hierarchic Sequential Prediction

We've seen syntactic structural knowledge constrains trees, but how do we guess trees from words?

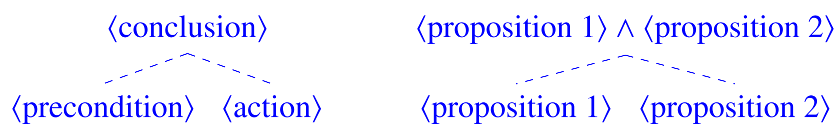
As an approximation, we base structural processing on domain-general complex event prediction.

We predict complex events in memory, using learned (generalized) rules of causation.

Contents

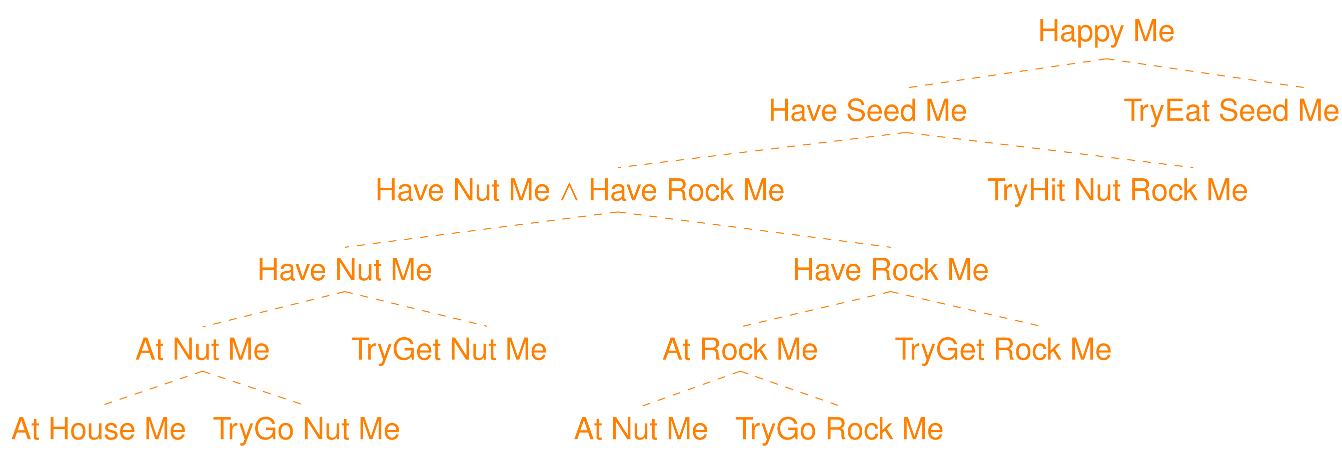
11.1 Complex events

We assume complex events are made of preconditions and actions, or conjoined propositions:



Events can contain hierarchies of subevents, especially complex plans (complex ideas).

Here's a complex event for breaking open a nut with a rock and eating the seed inside:



Sub-events are related to parent events by cued associations for 'cause' predications.

When similar (recognition) operations are nested inside other operations, a process is called **recursive**.

11.2 Event fragments

Humans and (some) animals can recognize and re-create complex hierarchic events.

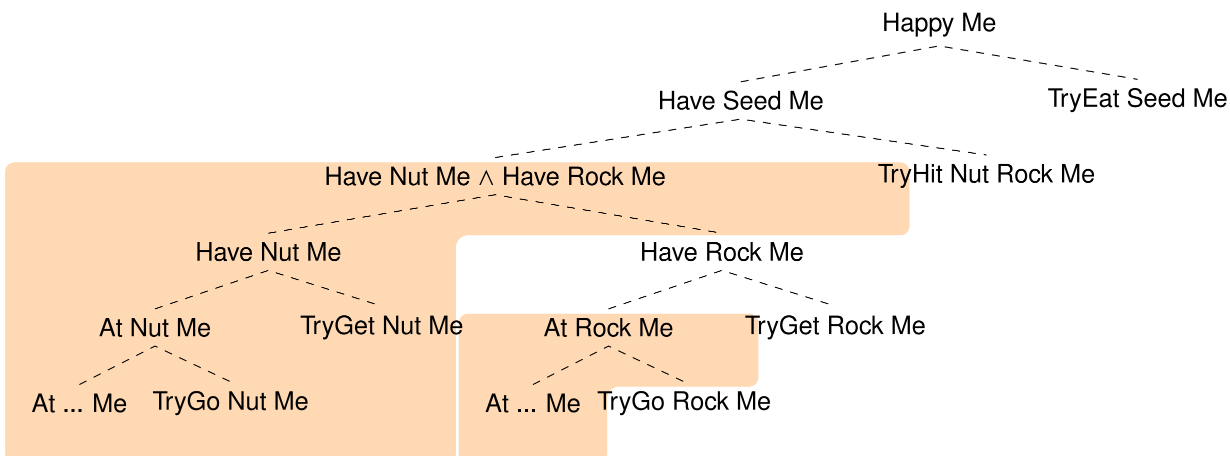
[Fuster, 1990, Botvinick, 2007]

Partial sequences of events can be grouped and stored as **event fragments** a/b, where:

- a is a **predicted outcome** 'apex' top-level event or sub-event,
- b is an **expected part** 'base' sub-event / observed event yet to come, which completes the outcome.

E.g. At Rock Me can be accounted as Have Rock Me / TryGet Rock Me.

Here's a set of event fragments recognized from observations in time order, up to TryGo Rock Me:



Near-complete sub-events can be chained together to save memory:

E.g. ... / Have Rock Me and At Rock Me form ... / TryGet Rock Me.

When a recent event fragment is completed, it can be added to an earlier event fragment.

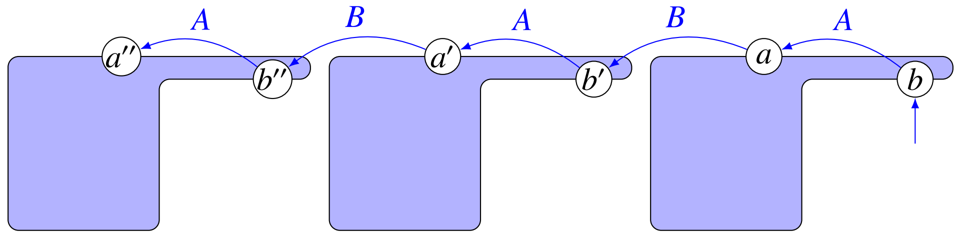
E.g. if At Rock Me is complete, it can satisfy Have Rock Me with TryGet Rock Me expected.

Uncertainty about events may be modeled using superposed activation vectors, described earlier.

11.3 Recognition Model [Johnson-Laird, 1983]

This model maintains a sequence of event fragments accessible from a most recent expectation b:

- Cued associations 'A' directly link individual expectations b to supported outcomes a.
- Cued associations 'B' directly link individual outcomes a to preceding expectations b.



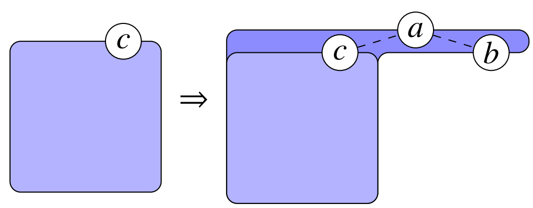
E.g. a'/b' is ... / Have Rock Me, a/b is At Rock Me / TryGo Rock Me.

This sequence is called a **store**. It is accessible from a **focus of attention** (straight arrow).

(You can think of the focus of attention as just what is currently active in the cortex.)

Crucially, this store can only be a few elements long before interference causes trouble.

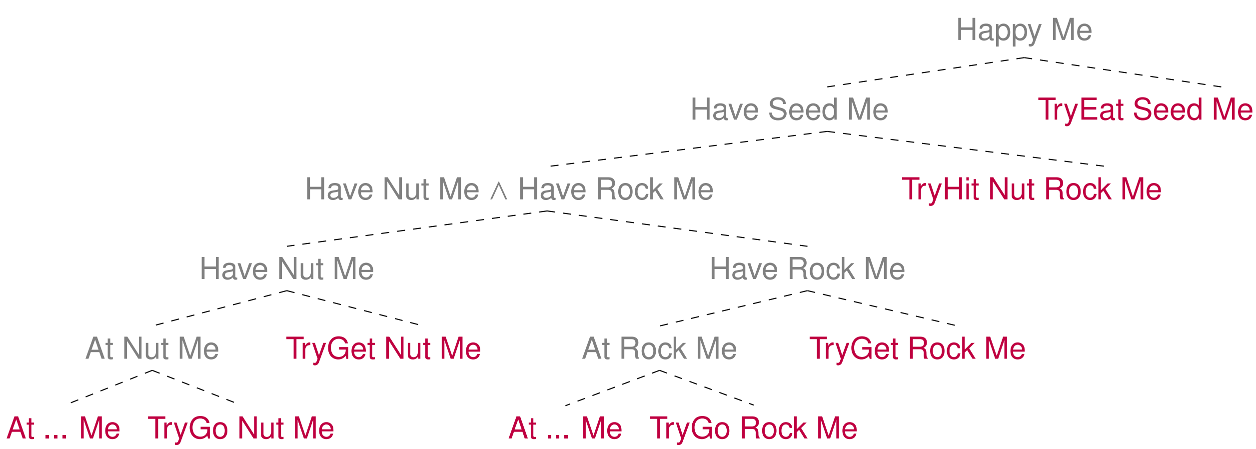
The model also assumes a set of learned **prediction rules**:



E.g. Have Rock Me (a) is composed of At Rock Me (c) followed by TryGet Rock Me (b).

Here, a, b, and c might be connected by a 'cause' elementary predication (dashed lines).

First, distinguish terminal (simple, observed) and nonterminal (complex, hidden) events:



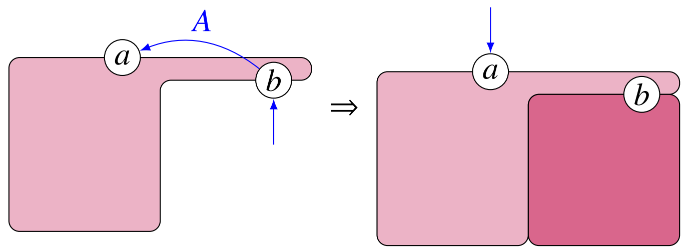
Note: in a binary-branching structure there are equal numbers of terminal and nonterminal events.

We predict this structure by guessing one terminal and one nonterminal branch at every observation.

Complex ideas can now be assembled by connecting observed events to event fragments...

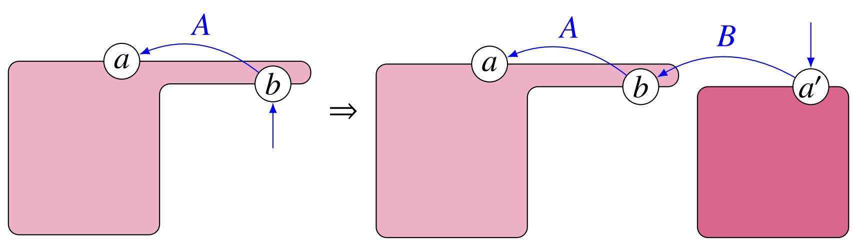
- **Terminal** decision (add observed event and connect to existing event fragment, or don't):

  Yes-match result (set current prediction):



  (Note that this replaces an event fragment with a complete event in associative memory.)

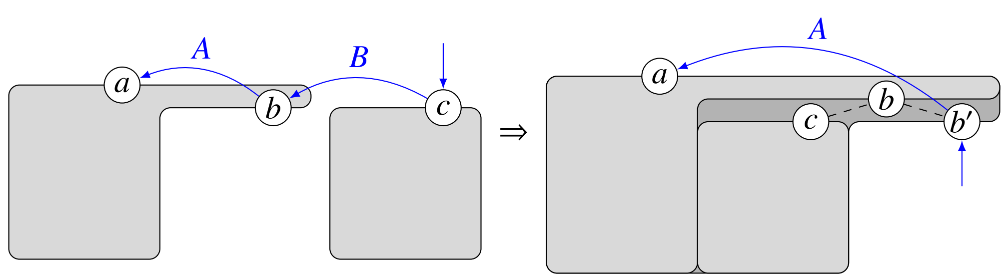
  No-match result (check types, store cued association from a' to b, set current prediction):



  (Note that this just adds a complete event to associative memory.)

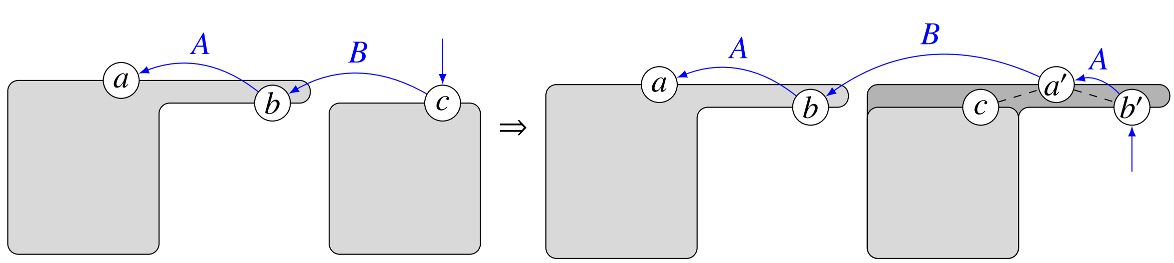
- **Nonterminal** decision (apply prediction rule and connect resulting event fragment, or don't):

  Yes-match result (check types, apply rule, store cued association from b' to a):



  (Note that this replaces an event fragment and a complete event with an event fragment.)

  No-match result (apply rule, store cued association from b' to a' and a' to b:



  (Note that this replaces a complete event with an event fragment in associative memory.)

Matching can be implemented in simple neural networks, generalized by procedural learning.

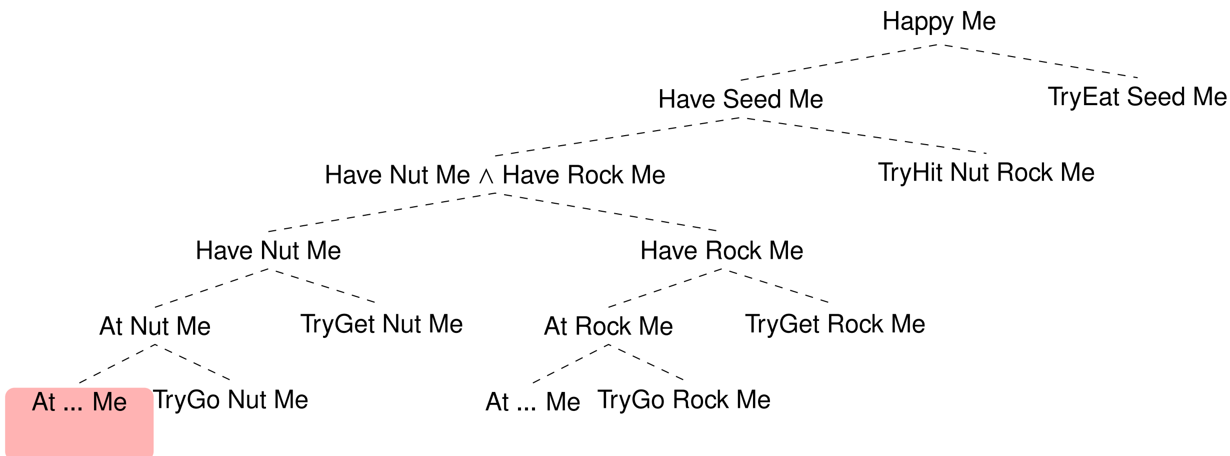
These operations can recognize any branching event structure using a minimum amount of memory.

11.4 Example recognition by hierarchic sequential prediction

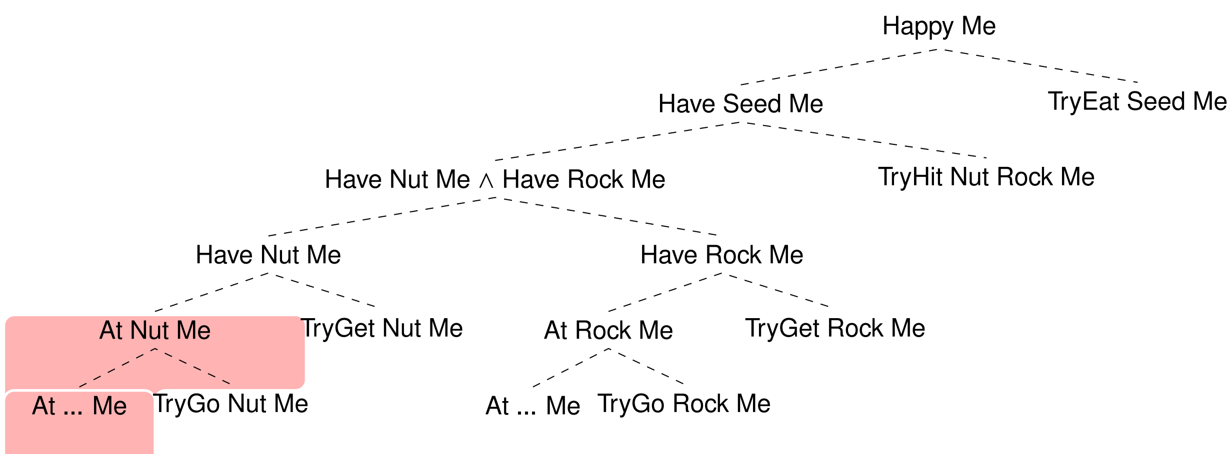
Here is an example of recognizing a complex plan from observations.

The events and event fragments will be drawn onto the phrase structure tree as they are recognized.

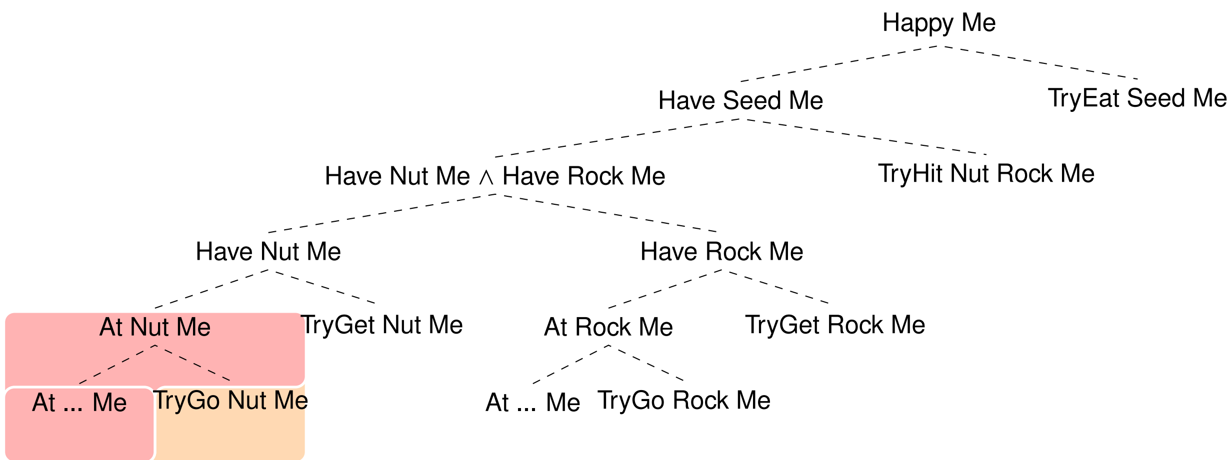
Terminal: start w. observation of At Me (wherever), don't match it to expectation of complete plan:



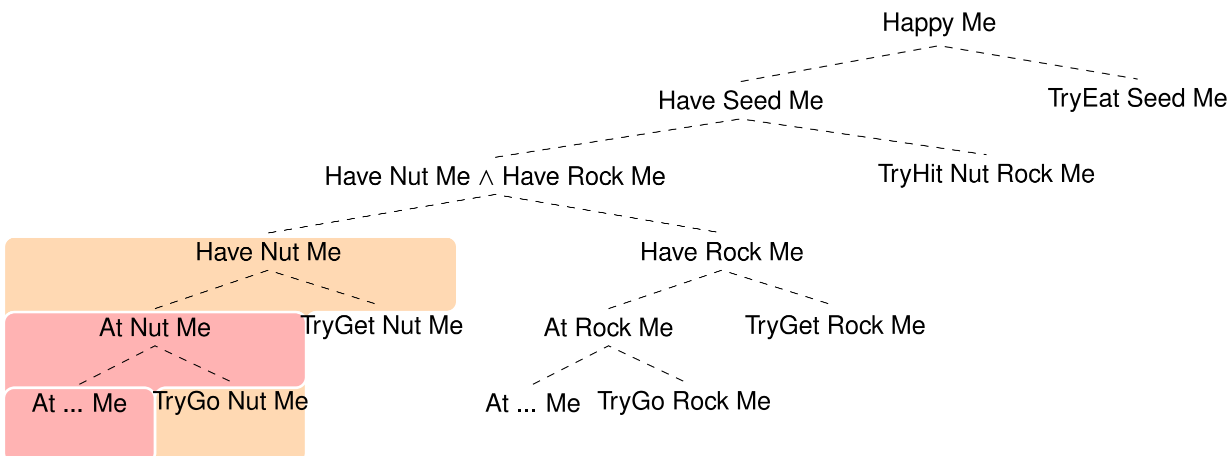
Nonterminal: expect TryGo Nut Me, outcome At Nut Me, don't match expectation of complete plan:



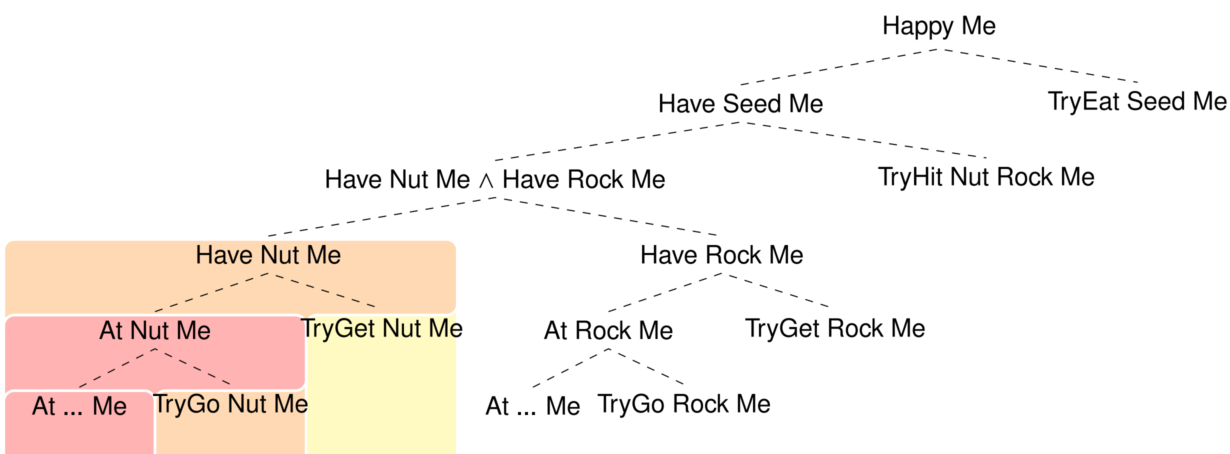
Terminal: observe TryGo Nut Me, match to expectation of TryGo Nut Me, making complete event:



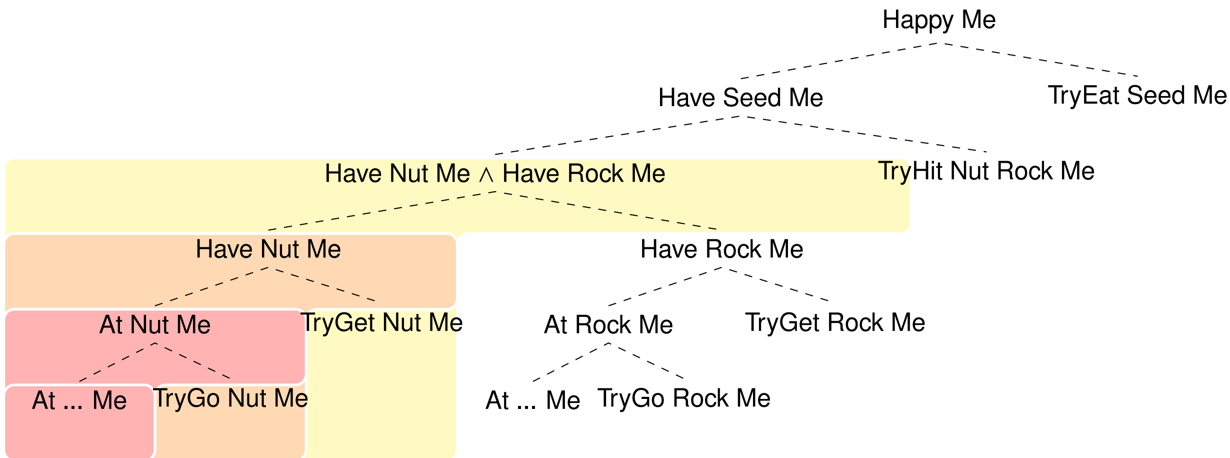
Nonterminal: expect TryGet Nut Me, outcome Have Nut Me, don't match to the completed plan:



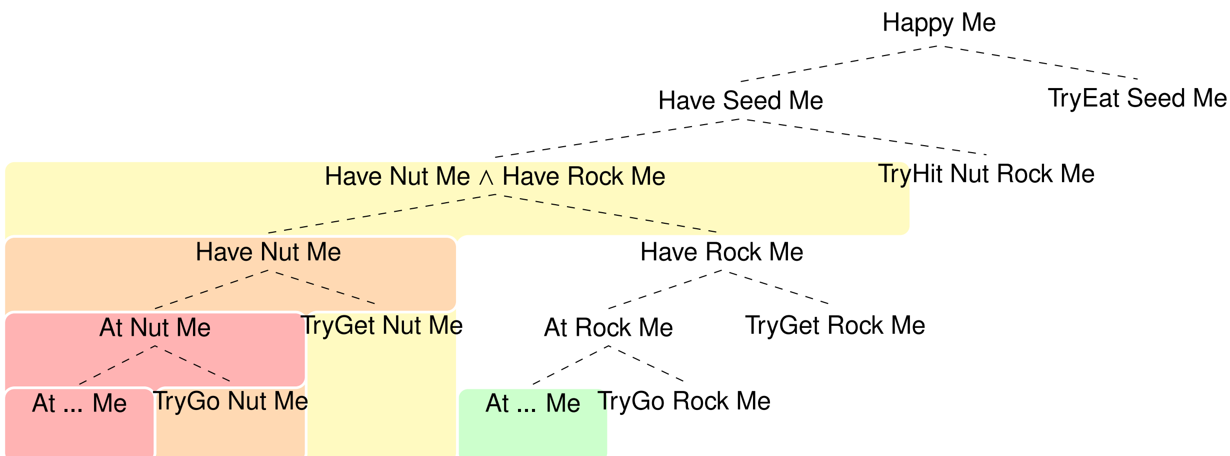
Term: observe TryGet Nut Me, match expectation of TryGet Nut Me in previous event fragment:



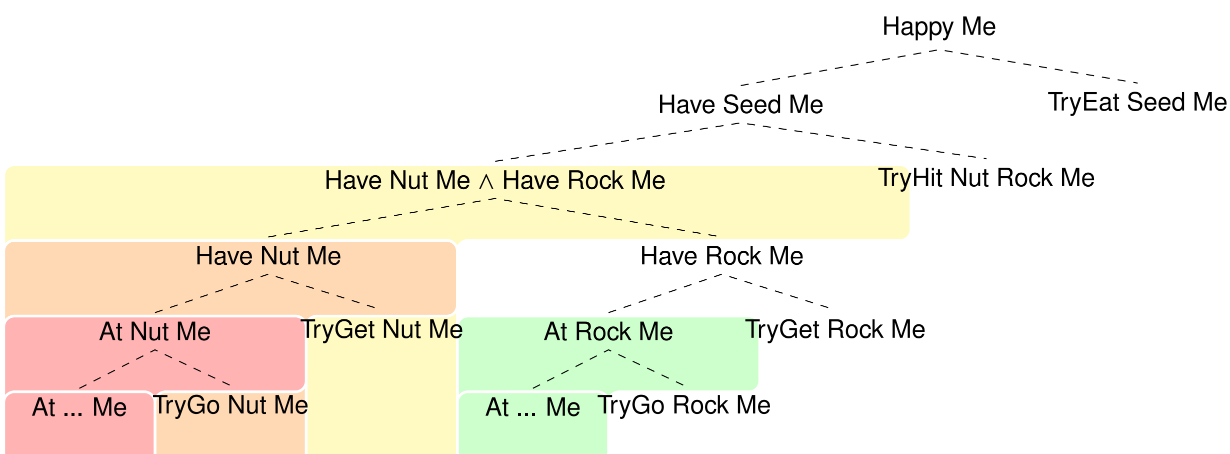
Nonterm: expect Have Rock Me, outcome Have Nut Me ∧ Have Rock Me, don't match as complete:



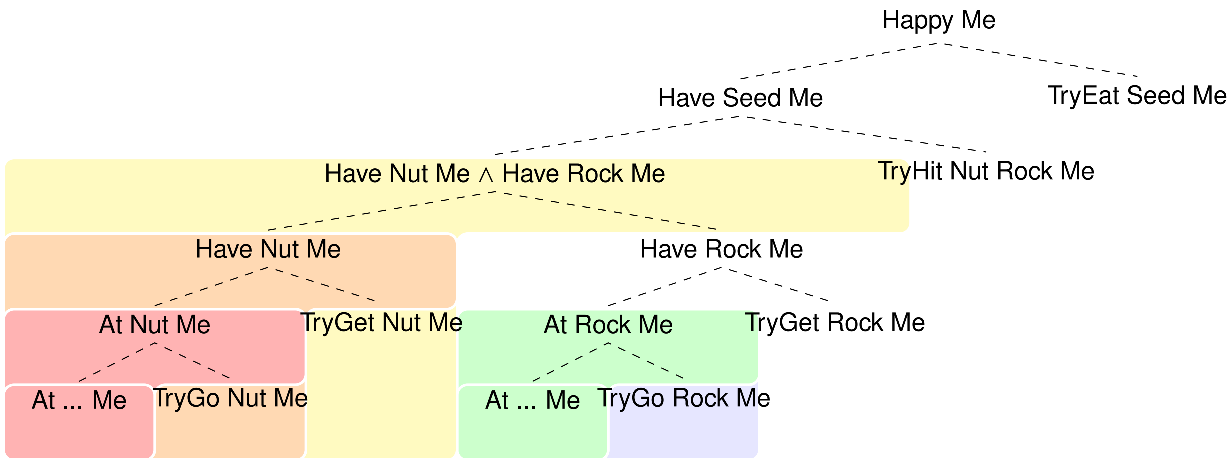
Term: observe At Me (wherever), don't match expectation Have Rock Me, creating separate event:



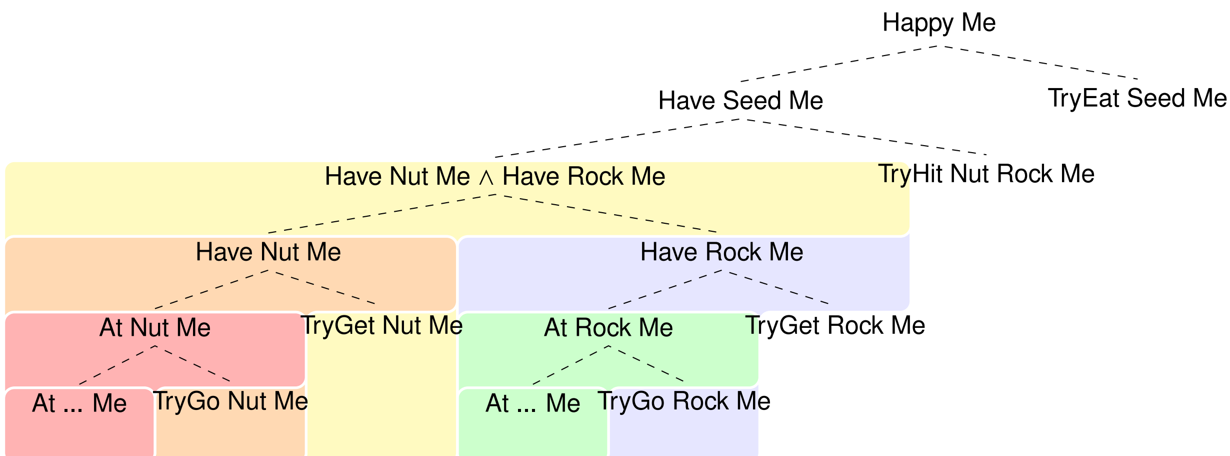
Nonterm: expect TryGo Rock Me, outcome At Rock Me, don't match expectation of Have Rock Me:



Term: observe TryGo Rock Me, match to expectation of TryGo Rock Me, forming complete event:



Nonterm: expect TryGet Rock Me, match outcome Have Rock Me, forming single event fragment:



Practice 11.1:

Continue the process for one more terminal and non-terminal decision.

Draw the event fragments that exist immediately after observing TryGet Rock Me. Specifically:

1. What will be the terminal decision result, and what event will exist afterward?
2. What will be the non-terminal decision result, and what fragment will exist afterward?

References

[Botvinick, 2007] Botvinick, M. (2007). Multilevel structure in behavior and in the brain: a computational model of Fuster's hierarchy. *Philosophical Transactions of the Royal Society, Series B: Biological Sciences*, 362, 1615–1626.

[Fuster, 1990] Fuster, J. M. (1990). Behavioral electrophysiology of the prefrontal cortex of the primate. *Progress in Brain Research*, 85, 313–324.

[Johnson-Laird, 1983] Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA, USA: Harvard University Press.