

Epistemic Semantics in Guarded String Models

Anonymous SCiL submission

Abstract

Constructive and computable multi-agent epistemic possible worlds models are defined, where possible worlds models are guarded string models in an epistemic extension of Kleene Algebra with Tests. The account is framed as a formal language Epik (Epistemic KAT) for defining such models. The language is implemented by translation into the finite state calculus, and alternatively by modeling propositions as lazy lists in Haskell. The syntax-semantics interface for a fragment of English is defined by a categorial grammar.

1 Introduction and Related Work

Linguistic semantics in the Montague tradition proceeds by assigning propositional *semantic values* to disambiguated sentences of a natural language. A proposition is a set or class of *possible worlds*. These are often assumed to be things with the same nature and complexity as the world we occupy (Lewis, 1986). But alternatively, one can work with small idealized models, in order to illustrate and test ideas. The point of this paper is to scale up toy or idealized models to countable sets of worlds, and to constructive and computable modeling of epistemic alternatives for agents. We describe a systematic way of defining such models, and illustrate how to apply them in natural language semantics. The focus is on epistemic semantics and clausal embedding. The fundamental move is to identify possible worlds with strings of primitive events, so that propositions are sets of strings. An advantage in this is that it allows for a mathematical description of an algebra of propositions, coupled with a computational representation using either lazy lists of strings, or finite state machines that describe sets of strings.

The approach taken here synthesizes five antecedents in a certain way. John McCarthy's *Situation*

Calculus is the source of the idea of constructing possible worlds as event sequences (McCarthy, 1963; Reiter, 2001). The algebraic theory of *Kleene Algebra with Tests* characterizes algebras with elements corresponding to propositions and event types in our application (Kozen, 2001). *Action models* in dynamic epistemic semantics introduce the technique of constructing epistemic models from primitive alternative relations on events, in order to capture the epistemic consequences of perceptual and communicative events (Baltag et al., 1999). Literature on *finite state methods in linguistic semantics* has used event strings and sets of event strings to theorize about tense and aspect in natural language semantics (Fernando, 2004, 2007; Carlson, 2009) and to express intensions (Fernando, 2017). Work on *finite state intensional semantics* has investigated how to do the semantics of intensional complementation in a setting where compositional semantics is expressed in a finite state calculus (Rooth, 2017; Collard, 2018).

A running example of an event-sequence model is *The Concealed Coin*. Amy and Bob are seated at a table. There is a coin on the table under a cup, heads up. The coin could be heads-up *H* or tails-up *T*, and neither agent knows which it is. This initial situation is possible world w_1 . Two additional worlds w_2 and w_3 are defined by sequencing events after the initial state, with events interpreted as in (1). The truth values for English sentences shown in (3) are observed, where 0 stands for falsity and 1 for truth.

- (1) a_1 Amy peeks at *H*, by tipping the cup. Bob sees she's peeking, but not what she sees.
- b_1 Bob peeks at *H*.
- a_0 Amy peeks at *T*.
- b_0 Bob peeks at *T*.
- a_{01} Amy secretly turns the coin from *T* to *H*. She knows she turned the coin over, but

not which side was face up. Bob thinks nothing happened.

a_{10} Amy secretly turns the coin from H to T .

a_{01} Bob secretly turns the coin from T to H .

b_{10} Bob secretly turns the coin from H to T .

$$(2) \quad w_2 = w_1 a_1 \quad w_3 = w_2 b_1 \quad w_4 = w_3 a_{10} b_{01} b_1$$

(3) $w_1 \quad w_2 \quad w_3 \quad w_4$ Sentence

0 1 1 0 Amy knows it's heads.

0 0 1 1 Bob knows it's heads.

0 0 1 0 Bob knows Amy
knows it's heads.

0 1 1 0 Bob knows Amy
knows whether it's
heads or tails.

The events come with pre-conditions and post-conditions. Amy can turn the coin from heads to tails only if the coin is heads-up, so a_{10} has the pre-condition of the coin being heads up. Once she turns the coin over, tails must be face-up, so a_{10} has the post-condition of the coin being tails-up. Let h be the Boolean proposition that the coin is heads up and t be the Boolean proposition that the coin is tails-up. Then pre- and post-conditions can be described by Boolean formulas, with h being the pre-condition of a_{10} and a_{01} being the post-condition. This is expressed using an operator “ $:$ ” (read “and next”) that pairs Boolean formulas. The formula $h : t$ describes a_{10} (Amy turning the coin from heads to tails) as happening only in an h state, and concluding in a not- h state. Events don’t have to change state: the event a_0 (Amy peeking at tails) can happen only in a t state, and does not change the state ($t : t$).

However a coin cannot be showing both heads and tails! Currently the precondition h of a_{10} only says that heads must be showing, and says nothing about the fact that tails must be face-down, indicated by the formula \bar{t} . We will further restrict the feasible conditions for our actions by restricting the space of valuations for our formulae.

A sequence such as $\bar{h}t$ can be viewed both as a formula and as a valuation of primitive propositions, which we use to describe world state. The primitives are listed in fixed order, and left unmarked (indicating true) or marked with the overbar (indicating false). Since a coin is heads or tails but not both, we want to allow the valuations $h\bar{t}$ and $\bar{h}t$, and disallow ht and $\bar{h}\bar{t}$. This is enforced by a *state formula*, which is a Boolean formula, in this case the one given on the second line of (4). Where B is

state formulas	$(a \in B)$	150
$\rho, \sigma, \varphi ::= a \mid 0 \mid 1 \mid \rho + \sigma \mid \rho \sigma \mid \bar{\rho}$		151
effect formulas		152
$\zeta, \eta ::= \rho : \sigma \mid \zeta + \eta \mid \zeta \& \eta \mid \bar{\zeta}$		153
$[\![\rho : \sigma]\!]^\varphi = A_B^{\rho\varphi} \times A_B^{\sigma\varphi}$		154
$[\![\zeta + \eta]\!]^\varphi = [\![\zeta]\!]^\varphi \cup [\![\eta]\!]^\varphi$		155
$[\![\zeta \& \eta]\!]^\varphi = [\![\zeta]\!]^\varphi \cap [\![\eta]\!]^\varphi$		156
$[\![\bar{\zeta}]\!]^\varphi = A_B^\varphi \times A_B^\varphi \setminus [\![\zeta]\!]^\varphi$		157

Figure 1: Syntax of state formulas and syntax and semantics of effect formulas. Effect formulas denote relations between atoms. In a state formula, juxtaposition $\rho\sigma$ is conjunction.

a set of atomic tests and ϕ is a state constraint over \mathbf{B} , $\mathcal{A}_{\mathbf{B}}^\phi$ is the set of valuations of \mathbf{B} that make formula ϕ true. Valuations are called atoms, because they correspond to the atoms of a Boolean algebra of tests (Kozen, 2001).

Formulas that describe pre- and post-conditions are *effect formulas*. They are interpreted as defining relations between atoms, as defined in Figure 1. The atoms they relate are constrained by the state formula as well. For the heads-tails example, let the state formula and the effect formula for a_1 (Amy peeking at heads) be as specified in (4). Then \mathcal{A}_B^φ and the relation on atoms for the event a_1 are as given at the bottom in (4).

(4)	B	$\{h, t\}$
	state formula φ	$h\bar{t} + \bar{h}t$
	effect formula ζ for a_1	$h : h$
	\mathcal{A}_B^φ	$\{\bar{h}\bar{t}, \bar{h}t\}$
	$[\zeta]^\varphi$	$\{\langle h\bar{t}, \bar{h}t \rangle\}$

2 Epistemic guarded string models

Epik is a specification language for possible worlds models that includes declarations of events and states, state formulas, effect formulas, and additional information. Figure 2 shows an Epik program that describes a possible worlds model for two agents with information about one coin, events of the agents semi-privately looking at the coin, and events of secretly turning the coin. The line beginning with `state` enumerates B. The line beginning with `restrict` gives the state formula. The lines beginning with `event` declare events and their effect formulas. Finally the lines beginning with `agent` define *event alternative* relations for agents. Each clause with an arrow has a single event symbol on the left, and a disjunction of alternative events on the right of the arrow. The in-

```

200
201 state h t      agent amy
202 restrict h!t    o1 -> o1
203           + t!h  o0 -> o0
204 event o1 h:h   a1 -> a1
205 event o0 t:t   a0 -> a0
206 event a1 h:h   b1 -> b1 + b0
207 event a0 t:t   b0 -> b1 + b0
208 event b1 h:h   a10 -> a10 + a01
209 event b0 t:t   a01 -> a10 + a01
210 event a10 h:t  b10 -> o0 + o1
211 event a01 t:h  b01 -> o0 + o1
212 event b10 h:t
213 event b01 t:h  agent bob
214           <sim. swap a and b>

```

Figure 2: Epik program describing a possible-worlds event sequence model for two agents with information about one coin, and events of the agents semi-privately looking at the coin.

terpretation of Amy’s alternatives for b_1 (Bob peeks at heads), is that when b_1 happens, for Amy either b_1 or b_0 (Bob peeks at tails) could be happening. Her alternatives for a_{01} and a_{10} (she turns the coin over) are a_{10} and a_{01} , indicating that she doesn’t know, *a priori*, whether she’s turning the coin from H to T or from T to H . Similarly, Bob secretly turns the coin over, in b_{10} or b_{01} , she doesn’t know anything has happened, so her alternatives are the no-operation events o_1 and o_0 for heads-worlds and tails-worlds respectively. Bob’s alternatives are the same, *mutatis mutandi*.

This paper focuses on defining a concrete possible worlds model from an Epik specification. The models are an extension of guarded-string models for Kleene Algebra with Tests (KAT). This is an algebraic theory that has model classes including guarded string models, relational models, finite models, and matrix models. Our definitions and notation follow (Kozen, 2001). We add syntax and semantics to cover multi-agent epistemic semantics.

Guarded strings over a finite alphabet P are like ordinary strings, but with atoms over a set B alternating with the symbols from P . In the algebra described by Figure 2, P is the set of events $\{a_1, a_0, b_1, b_0, a_{10}, a_{01}, b_{10}, b_{01}\}$, and B is $\{h, t\}$.

In the coin example, as we already saw in (4), \mathcal{A}_B^φ is $\{h\bar{t}, \bar{h}t\}$, for which we use the shorthand $\{H, T\}$. A guarded string over P and B is a strings of events from P , alternating with atoms over B , and beginning and ending with atoms. In this construction, $w_1 = H$, $w_2 = Ha_1H$, $w_3 = Ha_1Hb_1H$, and $w_4 = Ha_1Hb_1Ha_{10}Tb_{01}H$.

The discussion of (2) mentioned building worlds by incrementing worlds with events. This is ac-

complished in guarded string models with fusion product \diamond , a partial operation that combines two guarded strings, subject to the condition that the atom at the end of the first argument is identical to the atom at the start of the second one. (5) gives some examples.

$$(5) \quad H b_1 H \diamond H a_1 H = H b_1 H a_1 H \\ T b_{01} H \diamond T a_1 T = \text{undefined}$$

Rather than individual guarded strings, elements of a guarded string model for KAT are sets of guarded strings. In our application, these elements have the interpretation of propositions, which are sets of possible worlds. In a free guarded string model for KAT, any event can be adjacent to any atom in a guarded string that is an element of the underlying set for the algebra. We instead impose the constraints coming from the state and effect formulas. (6) defines the well-formed guarded strings determined by an Epik specification. Condition (i) says that each atom is consistent with the state constraint, and condition (ii) says that each constituent token event $\alpha_i e_i \alpha_{i+1}$ is consistent with the effect constraint on e_i .¹

$$(6) \quad \text{Given } P, B, \text{ a state formula } \varphi, \text{ and an effect formula } \zeta_e \text{ for each event } e \text{ in } P, \\ \alpha_0 e_0 \dots e_n \alpha_{n+1} \text{ is well-formed iff}$$

- (i) $\alpha_i \in \mathcal{A}_B^\varphi$ ($0 \leq i \leq n$), and
- (ii) $\langle \alpha_i, \alpha_{i+1} \rangle \in [\zeta_e]^\varphi$, ($0 \leq i \leq n$).

Well-formed guarded strings have the interpretation of worlds in the application to natural-language semantics. The set of possible worlds in the Kripke frame determined by an Epik specification is the set of well-formed guarded strings. At this point, we could say that any set of worlds is a proposition, so that the set of propositions is the power set of the set of worlds (Montague and Thomason, 1975; Gallin, 1975). We will instead define a more restrictive set of propositions corresponding to the regular sets of strings. This is deferred to the next section. Certain sets of well-formed guarded strings have the additional interpretation of event types. An event-type

¹An alternative is to define equations such as $\bar{\phi} = 0$ (from the state formula ϕ) and $a_1 = ha_1h$ (from the effect formula $h : h$ for event a_1), and construct a quotient algebra from the equivalence relation generated by these equations. This results in equating sets of guarded strings in the free algebra that differ by guarded strings that are ill-formed according to the state and effect formulas. In the development in the text, we instead use a set of guarded strings that are well-formed according to the state and effect formulas as the representative of the equivalence class.

300 is something that can “happen” in different worlds.
 301 For example, a_1 has the event type $\{H a_1 H\}$, and
 302 a_0 has the event type $\{T a_0 T\}$.

303 The construction so far defines a set of worlds
 304 from an Epik specification. Normally the set is
 305 countably infinite, though some choices of effect
 306 formulas can result in a finite set of worlds. The
 307 next step is to define an alternative relation R_a
 308 on worlds for each agent a . This will result in a
 309 Kripke frame $\langle W, R_1, \dots, R_n \rangle$ consisting of a set
 310 of worlds, and a world-alternative relation for each
 311 agent (Kripke, 1963). An Epik specification defines
 312 an alternative relation on bare events for each agent
 313 a , which we notate as R_a . This should be lifted
 314 to a relation \hat{R}_a on worlds. The basic idea is that
 315 when a world w is incremented with an event e , in
 316 the resulting world $w \diamond e$, epistemic alternatives for
 317 agent a are of the form $w' \diamond e'$, where w' is an alter-
 318 native to for a in w , and e' is an event-alternative
 319 to e for a .² This needs to be implemented in a way
 320 that takes account of pre- and post-conditions for
 321 events. For this, our approach is to refer the defini-
 322 tion of well-formed guarded strings. (7) defines a
 323 relation on worlds from a relation on bare events.

- 324 (7) Let W be a set of guarded strings over
 325 events P and primitive tests B , and R be
 326 a relation on P . The corresponding relation
 327 \hat{R} on W holds between a guarded string
 328 $\alpha_0 e_0 \dots e_n \alpha_{n+1}$ in W and a guarded string
 329 q iff q is an element of W and is of the
 330 form $\alpha'_0 e'_0 \dots e'_n \alpha'_{n+1}$, where for $0 \leq n$,
 331 $\langle e_i, e'_i \rangle \in R$.

332 This requires that in an alternative world, each
 333 constituent event e'_i is an alternative to the event e_i
 334 in the base world. Compatibilities between events
 335 in the alternative world are enforced by the require-
 336 ment that the alternative world is an element of W ,
 337 so that state and effect formulas are enforced.

338 Consider a scenario like the one from Figure 1,
 339 but with an additional agent Cal. The base world
 340 $Tb_0 Tc_0 T$ is one where the coin is tails, and first

341 ²In this it is important that the event-alternative relation
 342 for an agent is constant across worlds. We anticipate that the
 343 definition given here produces results equivalent to what is
 344 found in literature on event alternatives in dynamic epistemic
 345 semantics, though we have not verified this. That literature
 346 primarily focuses on mapping an epistemic model for a single
 347 time and situation to another, and uses general first-order mod-
 348 els, rather than guarded string models. See Baltag et al. (1999),
 349 Van Ditmarsch et al. (2007), and articles in Van Ditmarsch
 et al. (2015). Previous literature is motivated by epistemic
 logic and AI planning, rather than computable possible worlds
 models in natural language semantics.

events $e \in P$	350
$p, q ::= e \mid \sigma \mid p + q \mid pq \mid p^* \mid \neg p \mid \diamond_a p$	351
$\square_a p \triangleq \neg \diamond_a \neg p$	352
$\bullet \triangleq \sum_{e \in P} e$	353
$p \wedge q \triangleq \neg(\neg p + \neg q)$	354
$p \rightarrow q \triangleq \neg p + q$	355

356 Figure 3: The language of Epik terms and key derived
 357 operators.

358 Bob looks at tails, and them Cal looks at tails. The
 359 first event b_0 has the alternatives b_0 and b_1 for Amy,
 360 and the second event c_0 has the alternatives c_0 and
 361 c_1 for Amy. This results in four combinations
 362 $b_0 c_0$, $b_0 c_1$, $b_1 c_0$, and $b_1 c_1$. But these are filtered
 363 by post- and pre-conditions of events in the alter-
 364 native world, so that the set of alternatives for Amy
 365 in $Tb_0 Tc_0 T$ is $\{Tb_0 Tc_0 T, Hb_1 Hc_1 H\}$, with two
 366 world-alternatives instead of four.

3 The logical language of Epistemic KAT

368 The standard language for Kleene algebra with tests
 369 has the signature $\langle K, +, \cdot, *, \bar{}, 0, 1 \rangle$ (Kozen, 2001).
 370 In a guarded string model for KAT, K is a set of
 371 sets of guarded strings, $+$ is set union, the opera-
 372 tion \cdot is fusion product raised to sets, $*$ is Kleene
 373 star, the operation $\bar{}$ is complement for tests, 0 is
 374 the empty set, and 1 is the set of atoms.³ To this
 375 we add a unary modal operation \diamond_a for each agent,
 376 and a unary complement operation \neg on elements
 377 of K . Intuitively, $\diamond_a p$ is the set of worlds where
 378 proposition p is epistemically possible for agent
 379 a . Propositional complement is included because
 380 natural languages have sentence negation. In addi-
 381 tion, universal box modalities are defined as duals
 382 of existential diamond modalities.

383 With modalities and propositional negation
 384 added, the signature of n -agent epistemic KAT is
 385 $\langle K, +, \cdot, *, \bar{}, 0, 1, \neg, \diamond_1 \dots \diamond_n \rangle$. Figure 3 defines the
 386 syntax of the language. Juxtaposition is used for
 387 product. Terms in this language are used to repre-
 388 sent the propositional semantic values of English
 389 sentences. (8) gives some examples. To explain the
 390 first one, \bullet as defined in Figure 3 is the disjunction
 391 of the primitive events. Since a world is a well-
 392 formed sequence of events, \bullet^* is the set of worlds.
 393 Multiplying by the state symbol h in the term $\bullet^* h$
 394 has the effect of conjoining h with the atom at the
 395 end of the world. So $\bullet^* h$ is the set of worlds where
 396 the coin ends heads-up.

397 ³0 has the dual role the identity for $+$ (union), and as False
 398 for operations on tests. 1 has the dual role of the identity for
 399 product (fusion product raised to sets), and True for tests.

- 400 (8) $\bullet^* t$ It's tails.
 401 $\bullet^* h$ It's heads.
 402 $\bullet^* h \wedge \square_a \bullet^* h$ Amy knows it's heads.
 403 $\square_b(\square_a \bullet^* t + \square_a \neg \bullet^* t)$
 404 Bob believes Amy knows whether it's tails.

406 xs *Standard Epistemic Modalities* Using our existential modal primitive \diamond_a , and the dual encodings
 407 of \square_a and \wedge , we can encode the standard modal
 408 operators expressing knowledge (\mathcal{K}_a) and belief
 409 (\mathcal{B}_a) as in (9).⁴

411 (9) BELIEF $\mathcal{B}_a p \triangleq \square_a p$
 412 KNOWLEDGE $\mathcal{K}_a p \triangleq p \wedge \mathcal{B}_a p$

414 Different types of reasoners (e.g. accurate,
 415 inaccurate, etc) are modeled using the event-
 416 alternatives in an Epik specification.⁵ The agents
 417 in Figure 1 do not always have reliable beliefs, be-
 418 cause of the possibility of secret turning.

419 *Guarded String Interpretation.* A term p of the
 420 logical language is interpreted as a set of guarded
 421 strings $\llbracket p \rrbracket^{B,P,\varphi,\zeta}$, where superscript captures
 422 dependence on an Epik specification. Figure 4 defines
 423 the interpretation. The interpretation $\llbracket 1 \rrbracket^{B,P,\varphi,\zeta}$ of
 424 the multiplicative identity 1 is the set of atoms
 425 that satisfy the state constraint φ . Where b is a
 426 primitive Boolean, $\llbracket b \rrbracket^{B,P,\varphi,\zeta}$ is the set of atoms
 427 that satisfy the state constraint and where b is true.
 428 Where e is a primitive event, $\llbracket e \rrbracket^{B,P,\varphi,\zeta}$ is the set of
 429 guarded strings that have the form of e flanked by
 430 compatible atoms, as determined by the event for-
 431 mula ζ_e . The product pq is interpreted with fusion
 432 product raised to sets of guarded strings. Kleene
 433 star is interpreted as the union of exponents (p^n is
 434 the n -times product of p with itself, with $p^0 = 1$).
 435 Propositional complement is complement relative
 436 to the set of worlds. The epistemic formula $\diamond_a p$
 437 is interpreted with Kripke semantics for epistemic
 438 modality, as the pre-image of p under the world-
 439 alternative relation \hat{R}_a .

440 Summing up, given an Epik specification
 441 B, P, φ, ζ , term p (as defined syntactically in
 442 Figure 3) is interpreted as a set of guarded
 443 strings $\llbracket p \rrbracket^{B,P,\varphi,\zeta}$. Let $K^{B,P,\varphi,\zeta}$ be the sets
 444 that are interpretations of terms. Then
 445 $\langle K^{B,P,\varphi,\zeta}, +, \cdot, *, \bar{,}, 0, 1, \neg, \diamond_{a_1}, \dots, \diamond_{a_n} \rangle$ is a concrete
 446 guarded string interpretation for the signature

447 ⁴Deeper analysis of the lexical semantics of *know* requires
 448 adding modeling of presupposition (Collard, 2018). The gram-
 449 mar fragment in Section 6 does not model the presupposition
 450 of *know*, except as an entailment.

5 See the discussion of modal axioms **T** and **D** below.

$\llbracket 0 \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \emptyset$	450
$\llbracket 1 \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \mathcal{A}_B^\varphi$	451
$\llbracket b \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \mathcal{A}_B^{b\varphi}$	452
$\llbracket \sigma \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \mathcal{A}_B^\varphi \setminus \llbracket \sigma \rrbracket^{B,P,\varphi,\zeta}$	453
$\llbracket e \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \{\alpha e \beta \alpha \zeta_e \beta\}$	454
$\llbracket p + q \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \llbracket p \rrbracket^{B,P,\varphi,\zeta} \cup \llbracket q \rrbracket^{B,P,\varphi,\zeta}$	455
$\llbracket pq \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \left\{ \begin{array}{l l} x \diamond y & x \in \llbracket p \rrbracket^{B,P,\varphi,\zeta} \\ & y \in \llbracket q \rrbracket^{B,P,\varphi,\zeta} \\ & x \diamond y \text{ is defined} \end{array} \right\}$	456
$\llbracket p^* \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \bigcup_{n \geq 0} \llbracket p^n \rrbracket^{B,P,\varphi,\zeta}$	457
$\llbracket \neg p \rrbracket^{B,P,\varphi,\zeta}$	$\triangleq \llbracket \bullet^* \rrbracket^{B,P,\varphi,\zeta} \setminus \llbracket p \rrbracket^{B,P,\varphi,\zeta}$	458
$\llbracket \diamond_a p \rrbracket$	$\triangleq \{x \exists y. x \hat{R}_a y \wedge y \in \llbracket p \rrbracket^{B,P,\varphi,\zeta}\}$	459

460 Figure 4: Interpretation of Epik terms as sets of
 461 guarded strings

462 of epistemic KAT, with operations as in Figure
 463 4 (e.g. the binary operation $+$ is union, and the
 464 unary operation \diamond_a is pre-image relative to \hat{R}_a).
 465 This provides a concrete n -agent Kripke frame
 466 $\langle \llbracket \bullet^* \rrbracket^{B,P,\varphi,\zeta}, \hat{R}_1, \dots, \hat{R}_n \rangle$.⁶ The frame consists of
 467 a set of worlds, and an epistemic-alternative re-
 468 lation for each agent. It is used as a target for
 469 natural-language interpretation in Section 6.

470 *Axiomatic Classification.* To situate our logic as
 471 a modal logic, consider the soundness of the stan-
 472 dard modal axioms given our semantics (Hughes
 473 et al., 1996). Some of these standard axioms (see
 474 (10)) hold all the time, i.e. they are *valid*, and the
 475 remaining axioms are valid when \hat{R}_a has a certain
 476 shape. The axioms in (11) have a nontrivial condi-
 477 tion validity condition.

- 478 (10) **N** If p is valid, then $\square_a p$ is valid
 479 **K** $\square_a(p \rightarrow q) \rightarrow \square_a p \rightarrow \square_a q$ is valid.
 480 (11) **T** $\square_a p \rightarrow p$ if $g \hat{R}_a g, \forall g$
 481 **D** $\square_a p \rightarrow \diamond_a p$ if $g \in \text{dom}(\hat{R}_a), \forall g$
 482 **4** $\square_a p \rightarrow \square_a \square_a p$ if \hat{R}_a idempotent

483 The condition on **4** is essentially just a restate-
 484 ment of the theorem in relational terms. The condi-
 485 tions on the remaining standard modal axioms, **B**
 486 and **5**, are of a similarly trivial flavor.

4 Translation into the finite state calculus

492 The finite state calculus is an algebra of regular sets
 493 of strings and regular relations between strings that
 494 was designed for use in computational phonol-
 495

496 ⁶The domain of the Kripke frame differs from the domain
 497 of the guarded string model, because the former is the set of
 498 worlds, while the latter is the set of propositions.

```

500      St
501          Atomic Tests such as 0110.
502      UnequalStPair
503          Sequence of two unequal tests such as 0110 0111.
504      define Wf0 ~[$ UnequalStPair];
505          String that doesn't contain a non-matching test pair.
506      define Squash St -> 0 || St _;
507          Rewrite relation deleting the second of two tests.
508      define Cn(X, Y)
509          [[[X Y] & Wf0] .o. Squash].l;
510          KAT product.
511      define Kpl(X)
512          [[[X+] & Wf0] .o. Squash].l;
513      define Kst(X) St | Kpl(X);
514          KAT Kleene plus and Kleene star. The Fst operation | is
515          union.

```

Figure 5: Definition in Fst of KAT product and star.

ogy and morphographemics (Kaplan and Kay, 1994; Beesley and Karttunen, 2003). Current implementations allow for the definition of functions on regular sets and relations (Hulden, 2009; Karttunen, 2010). Such definitions are used here to construct of a model for epistemic KAT inside the finite state calculus. We describe our translation from Epik terms to Fst programs here.

The space of worlds is a set of ordinary (as opposed to guarded) strings. Bit sequences (sequences of 0's and 1's) encode atoms, and as before, these alternate with event symbols to encode a world. In this construction, w_3 of the example is the string 10 a1 10 b1 10.

Terms in the finite state calculus are interpreted as sets of strings, or for relational terms, as relations between strings. Computationally, the sets and relations are represented by finite state acceptors. As used here, a program in the Fst language of the finite state calculus is a straight-line program that defines a sequence of constants naming sets, constants naming relations, and functions (defined as macros) mapping one or more regular sets or relations to a regular set or relation.

Translating the Epik terms 0, 1, b , and e are straightforward: we simply convert the atoms as previously described, decorate the events e with their compatible atoms. For example a_1 becomes an Fst term denoting $\{01a_101\}$. Fst has built-in operations of union ($|$), and intersection ($\&$), which define the sum and intersection operations in the guarded string algebra. Fst set difference ($-$) is used to define propositional complement as the difference between the set of worlds and the argument.

Defining KAT product using Fst's set-lifted string concatenation (denoted by juxtaposition XY) requires more care. Naively concatenating strings

```

define RelKpl(R)
  Squash.i .o. Wf0 .o. [R+] .o. Wf0 .o. Squash
  a  Relational Kleene plus in the string algebra
  b  Constrain domain and co-domain to contain
     no unmatched tests.
  c  Reduce doubled tests to a single
     test in the domain and co-domain.
define RelKst(R) [St .x. St] | Kpl(X);
  The Fst operation .x. is Cartesian product. R.i is the
  inverse of relation R.

```

Figure 6: Definition in Fst of the Kleene concatenation closure of a relation between guarded strings.

with atoms (Boolean vectors) at both ends doubles atoms at the juncture, and does not enforce the requisite atom equality. To implement KAT product, we define the binary operation Cn , which concatenates strings in the string algebra, removes strings with non-matching atoms, and then deletes the second of two atoms to create a set of well-formed guarded strings. See Figure 5. $Wf0$ is the set of ordinary strings that does not contain unequal pairs of atoms, as defined using Fst's containment operator $\$$. The $Squash$ relation uses Fst's rewrite notation to delete atoms (elements of St) that are preceded by another atom.⁷ This relation is applied via the relational composition ($.o.$) and codomain ($.l$) operators.

KAT Kleene plus is defined in a similar way using Kleene plus in the string algebra, with checks for equality of atoms and deletion of atoms. See Figure 5. KAT Kleene star is defined from KAT Kleene plus and the multiplicative identity, which is the set of well-formed atoms St .

It remains to define an epistemic alternative relation on worlds for each agent. The relevant information in Figure 2 is a relation between bare events for each agent. This determines a relation between bare events decorated with compatible atoms. In Fst, we use the closure of the concatenation product operation on relations to lift a relation on decorated events for an agent to the corresponding relation on worlds. The concatenation product RS of two relations R and S is the set of pairs of the form $\langle x_1x_2, y_1y_2 \rangle$, where $x_1 R y_1$, and $x_2 S y_2$. In Fst, $R+$ is the closure of relation R with respect to this operation. Figure 6 defines the corresponding oper-

⁷This is a non-equal length regular relation. The finite state calculus includes such relations, and they can be used with relation composition and relation domain and co-domain. They are restricted in that the complement and set difference for non-equal length relations is not defined. Epistemic alternative relations are equal-length relations.

ation on sets of guarded strings as encoded in Fst.⁸ The epistemic alternative relation on worlds for an agent is then defined as the KAT relational concatenation closure RelKst of the decorated-event alternative relation for the agent.

5 Bounded Lazy Interpretation

We also implement the semantics of Epik terms using lazy lists in Haskell, rather than the direct interpretation as sets. Using lists sidesteps checking the set invariant (elements are unique) for large sets, such as \bullet^* , and laziness allows us to delay computing these large sets until they are actually needed. To sidestep the infiniteness of models, we parameterize the interpretation function on a positive integer n and only produce guarded strings of length n or less.

The bounded interpretation into lists of strings is very similar to the unbounded interpretation into sets of strings, except for the bounds checking. The full details are shown in Figure ???. First note that when $n = 0$, the denotation is empty, denoted $[]$. Terms of the form 0 , 1 , e , and ψ have the same denotation as before, translated into a list (denoted $[S]$, for a set S). We compute atoms using BBDs, which concisely represent boolean functions (?).

We lift the remaining operators (except Kleene star) to their list equivalents: union becomes list append (written $++$); fusion product is lifted to lists instead of sets, negation is implemented using list difference (\setminus), and the modal operator lifts the alternative relation over lists of strings⁹. The only caveat to these direct interpretations is that we restrict the operators to have size $\leq n$, denoted as $l|_n$ for a list of guarded strings l .

The denotation of p^* uses the fact that p^* and $1 + p; p^*$ are equivalent, and decrements the size threshold on the recursive denotation of p^* by i , where i is the length of the longest (nonzero) string in the denotation of p , making sure to filter out guarded strings that are too long.

6 Syntax-semantics interface

English sentences are mapped to terms in the logical language via a semantically interpreted multi-

⁸Relation concatenation in Fst differs from relation composition (\circ) , and the closure under discussion here is the closure of the former rather than the latter.

⁹Figure 7 depicts this using the list comprehension notation, which is analogous to set builder notation, except that it is written using square brackets. Element order is evoked by the keyword `for`, rather than using the unordered `forall`.

$(p)_0^{B,P,\phi,\zeta} \triangleq []$	650
$(0)_n^{B,P,\phi,\zeta} \triangleq []$	651
$(1)_n^{B,P,\phi,\zeta} \triangleq [A_B^\varphi]$	652
$(e)_n^{B,P,\phi,\zeta} \triangleq [\alpha e \beta \mid \alpha \in \zeta_e \beta]$	653
$(b)_n^{B,P,\phi,\zeta} \triangleq [A_B^{b\psi}]$	654
$(p+q)_n^{B,P,\phi,\zeta} \triangleq (p)_n^{B,P,\phi,\zeta} ++ (q)_n^{B,P,\phi,\zeta}$	655
$(p; q)_n^{B,P,\phi,\zeta} \triangleq ((p)_n^{B,P,\phi,\zeta} \diamond (q)_n^{B,P,\phi,\zeta}) _n$	656
$(p^*)_n^{B,P,\phi,\zeta} \triangleq [] + ((p)_n^{B,P,\phi,\zeta} \diamond (p^*)_n^{B,P,\phi,\zeta}) _n$ where $i = \max\{1, \min\{ g \mid g \in (p)_n^{B,P,\phi,\zeta}\}\}$	657
$(\neg p)_n^{B,P,\phi,\zeta} \triangleq (\bullet^*)_n^{B,P,\phi,\zeta} \setminus (p)_n^{B,P,\phi,\zeta}$	658
$(\Diamond_a p)_n^{B,P,\phi,\zeta} \triangleq [g' \mid g' \hat{R}_a g, \text{ for } g \in (p)_n^{B,P,\phi,\zeta}]$	659

Figure 7: Bounded interpretation using lazy lists

modal categorial grammar, consisting of a lexicon of words, their categorial types, and interpretations in a logical lambda language. The grammar covers basic statives (*it’s heads*), *that*- and *whether*-complements of *know*, predicate and sentence negation, and predicate and sentence conjunction. Figure 8 gives illustrative lexical entries.¹⁰ The grammar and semantics are optimized for a simple fragment of English concerned with clausal complementation. The agent names *Amy* and *Bob* contribute the epistemic alternative relations for those agents, rather than individuals. The root verb *know* contributes existential modal force. The complementizers *that* and *whether* are the heads of their dominating clauses, and assemble an alternative relation, modal force, and proposition contributed by the complement. These complementizers introduce the dual via two negations, in order to express universal modal force.

Multimodal categories such as \backslash_D and \backslash_M are used to control the derivation—phrases with these top-level slashes can only combine syntactically as arguments. The semantic translations in the third column of Figure 8 use the logical language, incremented with lambda. The body of $\lambda x. \bullet^* h$, which is the semantic lexical entry for *heads*, is a term denoting the set of all worlds where the coin is heads, expressed as the set of all guarded strings that end with a Boolean valuation where the primitive proposition h (*it’s heads*) is true. The body of $\lambda p. \lambda R. \Diamond_R p$, which is the semantic lexical entry for *knows*, is a term denoting the pre-image

¹⁰Category symbols use Lambek/Bar-Hillel notation for slashes, so that $(d \backslash t)/(d \backslash_D t)$ combines with $d \backslash_D t$ on the right to give a value that combines with d on the left to give t . In the semantics, lambda abstractions with multiple parameters are written $\lambda x y. e$ rather than $\lambda x. \lambda y. e$. d is the category of *it*.

ITEM	TYPE	SEMANTICS
Amy	e	\hat{R}_a
Bob	e	\hat{R}_b
it	d	d
heads	$d \setminus DT$	$\lambda x. \bullet^* h$
tails	$d \setminus DT$	$\lambda x. \bullet^* t$
is	$(d \setminus t) / (d \setminus DT)$	$\lambda P x. P x$
knows	$(e \setminus t) / MT$	$\lambda p R. p \vee \diamond_{RP}$
believes	$(e \setminus t) / MT$	$\lambda p R. \diamond_{RP}$
that	$((e \setminus t) / MT) \setminus (e \setminus t) / t$	$\lambda p m R. \neg(m (\neg p) R)$
whether	$((e \setminus t) / MT) \setminus (e \setminus t) / t$	$\lambda p m R. \neg(m (\neg p) R) + \neg(m p R)$

Figure 8: Partial categorial grammar lexicon. The first column has a word form, the second column a categorial type, and third column a semantic translation in a logical language that extends the Epik term language with lambda.

of the world-alternative relation contributed by the subject. This is not the right semantics for *Amy knows that it's heads*, because it has an existential modality \diamond_{RP} , rather than an universal modality \square_{RP} . This is corrected by the complementizer *that* or *whether*, which introduces the dual.

Sentences are parsed with a chart parser for categorial grammar. The semantics for complex phrases are obtained by application of semantic translations, accompanied by beta reductions that eliminate all lambdas logical forms for clauses. In consequence, the semantic term translating a sentence is an Epik term. Such a term designates a set of possible words (guarded strings). By way of example, (12a) is an English sentence with conjunction and several levels of clausal embedding. Using the grammar and parser, the sentence is mapped to the term in (12b). (12c) shows a simplified logical from constructed form (12b) using logical equivalences. Either term is compiled in an implementation of the finite state calculus to a finite state machine with 10 states and 23 edges, which accepts a countably infinite set of worlds.¹¹ In this way the methodology “directly” represents the set of worlds denoted by (12a).

- (12) a. It's tails and Amy knows that Bob knows that Amy knows whether it's heads.
b. $\bullet^* t \wedge \neg(\neg(\neg(\neg(\neg(\neg(\neg(\neg(\bullet^* h \vee \diamond_a \neg \bullet^* h) + \neg(\bullet^* h \wedge \diamond_a \bullet^* h)) + \diamond_b \neg(\neg(\neg \bullet^* h \vee \diamond_a \neg \bullet^* h) + \neg(\bullet^* h \wedge \diamond_a \bullet^* h)))) + \diamond_a \neg(\neg(\neg(\neg(\neg \bullet^* h \vee \diamond_a \neg \bullet^* h) + \neg(\bullet^* h \wedge \diamond_a \bullet^* h)))) + \neg(\bullet^* h \wedge \diamond_a \bullet^* h)) + \neg(\bullet^* h \wedge \diamond_a \bullet^* h)))$

¹¹Machine sizes need not be small, especially as the cardinality of B increases. A certain Epik model with fourteen primitive tests has the set of worlds represented by a finite state machine with 184794 states and 257881 edges.

$$\begin{aligned} & \diamond_a \bullet^* h)) + \diamond_b \neg(\neg(\neg \bullet^* h \vee \diamond_a \neg \bullet^* h) + \\ & \neg(\bullet^* h \wedge \diamond_a \bullet^* h)))) \\ c. & \bullet^* t \wedge \mathcal{K}_a(\mathcal{K}_b(\mathcal{K}_a \bullet^* h + \mathcal{K}_a \neg \bullet^* h)) \\ d. & \text{Amy knows that it's tails.} \\ e. & \bullet^* t \wedge \neg \diamond_a \neg \bullet^* t \quad (\equiv \mathcal{K}_a \bullet^* t) \end{aligned}$$

Sentence (12d) is assigned the logical form (12e) by the grammar. Logical relations between propositions are checked in the finite state calculus by checking set-theoretic relations between sets of worlds. For instance entailment $p \rightarrow q$ is decided by checking in an interpreter for the finite state calculus whether $p - q$ is non-empty. In the model defined by Figure 2, the propositions (12b) and (12e) are independent. They are equivalent in a version without secret flipping.

7 Discussion

The methodology presented here is designed for use in research in linguistic semantics, and for education at the level of a second graduate course in formal semantics, covering intensionality. There are straightforward extensions to additional linguistic phenomena, such as tense and perfective aspect as in (13a), and the combination of metaphysical modality and prospective aspect in (13b).

- (13) a. Amy has learned that Bob had learned that it's heads.
b. Amy might learn that it's heads.

The model framework is a constructive branching-time framework with metaphysical modality and epistemic modality, which will be applicable in linguistic semantic research on combinations of tense, metaphysical modality, and epistemic complementation (Thomason, 1984; Abusch, 1998; Condoravdi, 2002). Connections with research on temporal constitution of events in a related formal setting remain to be explored (Fernando, 2004, 2007; Carlson, 2009).

The development here is concerned with defining concrete computable possible worlds models, and applying them in natural language semantics. Issues for further investigation are mathematical characterizations of epistemic KATs, e.g. sound and complete axioms, coalgebra, and decidability.

Open source code will be distributed in conjunction with the conference.

800 References

- 801 Dorit Abusch. 1998. Generalizing tense semantics for
802 future contexts. In *Events and grammar*, pages 13–
803 33. Springer.
- 804 Alexandru Baltag, Lawrence S Moss, and Slawomir
805 Solecki. 1999. The logic of public announcements,
806 common knowledge, and private suspicions.
- 807 Kenneth R Beesley and Lauri Karttunen. 2003. *Fi-*
808 *nite State Morphology*. Center for the Study of Lan-
809 guage and Inf.
- 810 Lauri Carlson. 2009. *Tense, Mood, Aspect, Diathesis*.
811 Book ms., University of Helsinki.
- 812 Jacob Collard. 2018. Finite state reasoning for presup-
813 position satisfaction. In *Proceedings of the First In-
814 ternational Workshop on Language Cognition and
815 Computational Models*, pages 53–62.
- 816 Cleo Condoravdi. 2002. Temporal interpretation of
817 modals: Modals for the present and for the past. *The
818 construction of meaning*, 59:88.
- 819 Tim Fernando. 2004. A finite-state approach to events
820 in natural language semantics. *Journal of Logic and
821 Computation*, 14(1):79–92.
- 822 Tim Fernando. 2007. Observing events and situations
823 in time. *Linguistics and Philosophy*, 30(5):527–550.
- 824 Tim Fernando. 2017. Intensions, types and finite-
825 state truthmaking. In *Modern Perspectives in Type-
826 Theoretical Semantics*, pages 223–243. Springer.
- 827 Daniel Gallin. 1975. *Intensional and Higher-order
828 Modal Logic: With Applications to Montague Se-
829 mantics*. North-Holland Publishing Company.
- 830 George Edward Hughes, Max J Cresswell, and
831 Mary Meyerhoff Cresswell. 1996. *A new introduc-
832 tion to modal logic*. Psychology Press.
- 833 Mans Hulden. 2009. Foma: a finite-state compiler and
834 library. In *Proceedings of the 12th Conference of
835 the European Chapter of the Association for Compu-
836 tational Linguistics: Demonstrations Session*, pages
837 29–32. Association for Computational Linguistics.
- 838 Ronald M Kaplan and Martin Kay. 1994. Regular mod-
839 els of phonological rule systems. *Computational lin-
840 guistics*, 20(3):331–378.
- 841 Lauri Karttunen. 2010. Update on finite state morphol-
842 ogy tools. *Ms., Xerox Palo Alto Research Center*.
- 843 Dexter Kozen. 2001. Automata on guarded strings and
844 applications. Technical report, Cornell University.
- 845 Saul Kripke. 1963. Semantical considerations on
846 modal logic. *Acta Philosophica Fennica*, 16:83–94.
- 847 David Lewis. 1986. *On the Plurality of Worlds*. Black-
848 well.
- 849 John McCarthy. 1963. Situations, actions, and causal
laws. Technical report, Stanford CS.
- 850 Richard Montague and Richmond H Thomason. 1975.
851 Formal philosophy. selected papers of richard mont-
852 ague.
- 853 Raymond Reiter. 2001. *Knowledge in Action: Logical
854 foundations for specifying and implementing dynam-
855 ical systems*. MIT press.
- 856 Mats Rooth. 2017. Finite state intensional seman-
857 tics. In *IWCS 2017-12th International Conference
858 on Computational Semantics-Long papers*.
- 859 Richmond H Thomason. 1984. Combinations of tense
860 and modality. In *Handbook of philosophical logic*,
861 pages 135–165. Springer.
- 862 Hans Van Ditmarsch, Wiebe van Der Hoek, and Barteld
863 Kooi. 2007. *Dynamic Epistemic Logic*, volume 337.
864 Springer Science & Business Media.
- 865 Hans Van Ditmarsch, Joseph Y Halpern, Wiebe van der
866 Hoek, and Barteld Pieter Kooi. 2015. *Handbook of
867 Epistemic Logic*. College Publications.
- 868 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899