Prior Probabilities of Allen Interval Relations over Finite Orders

Tim Fernando and Carl Vogel
ADAPT Centre, Computer Science Department, Trinity College Dublin, Ireland

Keywords: Allen Interval Relations, Probabilities, Events.

Abstract: The probability that intervals are related by a particular Allen relation is calculated relative to sample spaces Ω_n given by the number n of, in one case, points, and, in another, interval names. In both cases, worlds in the sample space are assumed equiprobable, and Allen relations are classified as short, medium and long, according to the number of shared borders.

1 INTRODUCTION

A useful basis for relating intervals are the 13 relations described in (Allen, 1983) and widely applied to temporal relations in text and beyond (Liu et al., 2018; Verhagen et al., 2009; Allen and Ferguson, 1994; Kamp and Reyle, 1993, among many others). The present work proceeds from the following question.

(Q) Given an Allen relation R, what is the probability that R relates intervals a and a', aRa'?

Let us understand (Q) as saying nothing about a and a', not even that they are distinct (equality being an Allen relation). As there are 13 Allen relations, $\frac{1}{13}$ is a plausible answer to (Q), under the principle of indifference (commonly ascribed to Laplace). But are Allen relations a matter of indifference when, for example, some Allen relations occur more often than others in the transitivity table of (Allen, 1983)? That table is a central tool in interval networks formed from nodes representing intervals, and arcs labelled by Allen relations that may hold between the intervals. We will return to the transitivity table below. For now, suffice it to observe that some care is in order when proposing a sample space of equiprobable outcomes (hereafter, worlds) against which to answer (Q).

It is natural to interpret (Q) as presupposing a linear order relative to which a and a' are intervals. To accommodate all Allen relations, let us assume there are at least 4 points in that linear order, and for simplicity, let us suppose it is finite — say, the usual order on the set

$$[n] \ := \ \{i \in \mathbb{Z} \mid 1 \leq i \leq n\}$$

of integers between 1 and n (inclusive). A pair (l,r) from the linear order

$$<_n \ := \ \{(l,r) \in [n] \times [n] \mid l < r\}$$

on $[n]$ defines the $<_n$-interval

$$(l,r] \ := \ \{i \in [n] \mid l < i \leq r\}$$

(with *l*eft border l and *r*ight border r, allowing $=$ with r but not l). Now, over the linear order $<_n$, the probability that aRa' becomes the probability that

$$(l,r] \ R \ (l',r']$$

for (l,r) and (l',r') drawn from $<_n$. Note that 1 is excluded from $(l,r]$ for all $l,r \in [n]$. To lift this restriction, it suffices to work with copies in $<_{n+1}$ given by mapping $i \in [n]$ to $i+1 \in [n+1]$. Similarly, the requirement that a $<_n$-interval be strictly bounded to the right can be imposed by passing to $<_{n-1}$ with $i > 1$ mapped to $i-1$. Without loss of generality, we identify $<_n$-intervals with $(l,r]$ for $l <_n r$.[1]

Over a sample space Ω_n given by a linear order on n points, probabilities for each Allen relation R are calculated in section 2, under the assumption that worlds in Ω_n are equiprobable. The probabilities queried by (Q) vary with n and depend on the extent to which the intervals share borders, given R. As n approaches infinity, 7 of the 13 Allen relations have vanishing probabilities, leaving each of the other 6 probability $\frac{1}{6}$.

But should Allen probabilities be assessed around the number n of points in the linear order? The guiding perspective behind (Allen, 1983) (and many other

[1]As for why an interval should be half open and half closed, some motivation from Leibniz's law is presented in section 4 below.

works such as (Hamblin, 1971)) is that intervals, not points, are basic, suggesting that n pertain to intervals, not points. We take up this suggestion in section 3, working with interval names (also known as events under, for example, the Russell-Wiener construction of temporal instants described in (Kamp and Reyle, 1993, page 667)). Calculating probabilities becomes more complex without, as far as we can tell, straying from the asymptotic behavior determined in section 2: at the limit $n \to \infty$, 7 of the 13 Allen relations have probability 0, while 6 have $\frac{1}{6}$ each.

So what? The main thrust of this work is not so much to calculate numbers but to uncover structure lurking behind Allen relations. Concrete examples of structure in natural language semantics are described in the passage below from (Kamp, 2013, page 11)

> when we interpret a piece of discourse — or a single sentence in the context in which it is being used — we build something like a model of the episode or situation described; and an important part of that model are its event structure, and the time structure that can be derived from that event structure by means of Russell's construction.

The event structure Kamp has in mind is "made up by those comparatively few events that figure in this discourse" (page 9). The aforementioned Russell construction turns the finitely many events mentioned in a (finite) discourse into a finite linear order of temporal instants (each instant being a certain set of events). This contrasts sharply with the continuum $\mathbb{R}$ with which "real" time is commonly identified (Kamp and Reyle, 1993, for example) or, for that matter, any unbounded linear order for the time periods of (Allen and Ferguson, 1994). Indeed, if an event is equipped with its past and future — or, in the terminology of (Freksa, 1992), an interval is represented by its semi-intervals — then the resulting time structure amounts to ordering the left and right borders l and r of events (Fernando, 2016, page 3635). The case of two events yields the Allen relations, which can be formulated naturally in terms of strings (Durand and Schwer, 2008). That formulation is recounted in Table 1 in section 2 below.

The appeal to left and right borders runs counter to the use of the transitivity table in (Allen, 1983), where borders are buried out of sight. That said, both sections 2 and 3 end with links to the transitivity table. A more serious issue is the assumption of equiprobable worlds, which we reconsider in section 4, after the nature of the sample spaces becomes clearer. That space is formed in section 3 out of strings that go well beyond pictures of Allen relations between two intervals. Throughout this paper, however, our focus is on answering the question (Q) against a finite temporal structure (given by a finite discourse).

2 PROBABILITIES OVER n ORDERED POINTS

Let $\mathcal{AR}$ be the set of 13 names

$$\text{b, bi, d, di, o, oi, m, mi, s, si, f, fi, e}$$

of Allen relations. For each $R \in \mathcal{AR}$, Table 1 pictures $(l,r]\ R\ (l',r']$ as a string $\mathfrak{s}_R$ of boxes arranged from left to right so that all borders in the same box are equal and are $<$ borders in boxes to the right (Durand and Schwer, 2008).

Table 1: Allen relations in strings, following Figure 4 of (Durand and Schwer, 2008).

$(l,r]\ R\ (l',r']$	$\mathfrak{s}_R$	R^{-1}	$\mathfrak{s}_{R^{-1}}$
$(l,r]$ b $(l',r']$	$\boxed{l}\boxed{r}\boxed{l'}\boxed{r'}$	bi	$\boxed{l'}\boxed{r'}\boxed{l}\boxed{r}$
$(l,r]$ d $(l',r']$	$\boxed{l'}\boxed{l}\boxed{r}\boxed{r'}$	di	$\boxed{l}\boxed{l'}\boxed{r'}\boxed{r}$
$(l,r]$ o $(l',r']$	$\boxed{l}\boxed{l'}\boxed{r}\boxed{r'}$	oi	$\boxed{l'}\boxed{l}\boxed{r'}\boxed{r}$
$(l,r]$ m $(l',r']$	$\boxed{l}\boxed{r,l'}\boxed{r'}$	mi	$\boxed{l'}\boxed{r',l}\boxed{r}$
$(l,r]$ s $(l',r']$	$\boxed{l,l'}\boxed{r}\boxed{r'}$	si	$\boxed{l,l'}\boxed{r'}\boxed{r}$
$(l,r]$ f $(l',r']$	$\boxed{l'}\boxed{l}\boxed{r,r'}$	fi	$\boxed{l}\boxed{l'}\boxed{r,r'}$
$(l,r]$ e $(l',r']$	$\boxed{l,l'}\boxed{r,r'}$	e	

For example, $\boxed{l}\boxed{r}\boxed{l'}\boxed{r'}$ depicts the ordering

$$l < r < l' < r' \text{ characteristic of } (l,r]\ \text{b}\ (l',r']$$

while $\boxed{l,l'}\boxed{r,r'}$ depicts the ordering

$$l = l' < r = r' \text{ characteristic of } (l,r]\ \text{e}\ (l',r'].$$

Each $R \in \mathcal{AR}$ can be classified as either long

$$\{R \in \mathcal{AR} \mid \text{length}(\mathfrak{s}_R) = 4\} \ = \ \{\text{b,d,o,bi,di,oi}\}$$

or medium

$$\{R \in \mathcal{AR} \mid \text{length}(\mathfrak{s}_R) = 3\} \ = \ \{\text{m,s,f,mi,si,fi}\}$$

or short

$$\{R \in \mathcal{AR} \mid \text{length}(\mathfrak{s}_R) = 2\} \ = \ \{\text{e}\}$$

according to the length of $\mathfrak{s}_R$, which also happens to be the cardinality of the set $\{l,l',r,r'\}$ when $(l,r]\ R\ (l',r']$. The probabilities assigned in this paper to each $R \in \mathcal{AR}$ will turn out to depend on whether R is long, medium or short.

More precisely, given an integer $n \geq 4$, let us agree an *n-world* is a function

$$f : \{x,y,x',y'\} \to [n]$$

assigning four distinct variables x, y, x', y' integers in $[n]$ such that

$$f(x) < f(y) \ \text{ and } \ f(x') < f(y').$$

For each $R \in \mathcal{AR}$, we say an n-world f *satisfies* R if

$$(f(x), f(y)] \ R \ (f(x'), f(y')].$$

Now comes a key observation.

Lemma 1. *Given an integer $n \geq 4$,*

(i) *the number of n-worlds satisfying* e *(equal) is*

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

(ii) *for each medium $R \in \mathcal{AR}$, the number of n-worlds satisfying R is*

$$\binom{n}{3} = \binom{n}{2}\frac{n-2}{3}$$

(iii) *for each long $R \in \mathcal{AR}$, the number of n-worlds satisfying R is*

$$\binom{n}{4} = \binom{n}{3}\frac{n-3}{4}.$$

Proof. Let im be the map from an n-world f to its image

$$im(f) = \{f(x), f(y), f(x'), f(y')\} \subseteq [n].$$

For each $R \in \mathcal{AR}$, let im_R be the restriction of im to n-worlds satisfying R. It suffices to observe that im_R is a bijection to subsets of $[n]$ of cardinality

$$\begin{cases} 4 & \text{if } R \text{ is long} \\ 3 & \text{if } R \text{ is medium} \\ 2 & \text{if } R \text{ is e.} \end{cases}$$

□

Let Ω_n be the set of n-worlds, and for each $R \in \mathcal{AR}$, let $p_n(R)$ be the fraction of Ω_n satisfying R

$$p_n(R) \ = \ \frac{\text{cardinality}(\{f \in \Omega_n \mid f \text{ satisfies } R\})}{\text{cardinality}(\Omega_n)}.$$

Representing the medium relations by meet, m, and long relations by before, b, we have from Lemma 1,

$$\frac{p_n(\text{m})}{p_n(\text{e})} = \frac{n-2}{3} \ \text{ and } \ \frac{p_n(\text{b})}{p_n(\text{m})} = \frac{n-3}{4}$$

which with

$$1 \ = \sum_{R \in \mathcal{AR}} p_n(R) \ = \ p_n(\text{e}) + 6p_n(\text{m}) + 6p_n(\text{b})$$

allows us to solve for $p_n(\text{e})$. A simpler alternative suggested by a referee is to use

$$\text{cardinality}(\Omega_n) \ = \ \binom{n}{2} \cdot \binom{n}{2}$$

(as Ω_n consists of all choices of pairs l, r and l', r' from $[n]$). Either way, we obtain

Theorem 2. *For $n \geq 4$ and $R, R' \in \mathcal{AR}$,*

$$p_n(R) = p_n(R') \ \text{ if } \ \mathit{length}(\mathfrak{s}_R) = \mathit{length}(\mathfrak{s}_{R'})$$

where the short relation e *(equal) has probability*

$$p_n(\text{e}) \ = \ \frac{2}{n(n-1)}$$

while medium relations have probabilities

$$p_n(\text{m}) \ = \ \frac{2(n-2)}{3n(n-1)}$$

and long relations have probabilities

$$p_n(\text{b}) \ = \ \frac{(n-3)(n-2)}{6n(n-1)}.$$

Corollary 3. *For $R \in \mathcal{AR}$,*

$$\lim_{n \to \infty} p_n(R) \ = \ \begin{cases} 0 & \text{if } R \text{ is short or medium} \\ \frac{1}{6} & \text{otherwise.} \end{cases}$$

To put Corollary 3 in context, the probabilities at the start are strikingly different, with e the most probable at $n = 4$, m catching up at $n = 5$, and b at $n = 6$ (and the most probable from $n \geq 8$).

Table 2: Some probabilities from Theorem 2.

n	$p_n(\text{e})$	$p_n(\text{m})$	$p_n(\text{b})$
4	1/6	1/9	1/36
5	1/10	1/10	1/20
6	1/15	4/45	1/15
8	1/28	1/14	5/56

Recall from the Introduction that 1 should be added to or subtracted from n to lift or impose bounds. At any rate, there is an arbitrariness in any choice of n that calls out for attention. Letting n approach $+\infty$ (as in Corollary 3) is an admittedly crude way to attend to this. A more sophisticated approach would build on a probability distribution on the lengths n — a direction not pursued below.

What is pursued is the short-medium-long classification of Allen relations, which we pause now to note is implicit in the transitivity table at the center of (Allen, 1983). That table maps a pair (R_1, R_2) of Allen relations to the set $t(R_1, R_2)$ of Allen relations R such that there are intervals i, j and k for which

$$iR_1 j \text{ and } jR_2 k \text{ and } iRk.$$

Let us define the *t-number of* an Allen relation R to be the sum

$$\#(R) \ := \sum_{R' \in \mathcal{AR}} \text{cardinality}(t(R, R'))$$

of the numbers of entries in the row for R, including the Allen relation of equality, e, omitted from the transitivity table in (Allen, 1983), which we incorporate into t as expected

$$t(R, \mathrm{e}) \ = \ t(\mathrm{e}, R) \ = \ \{R\} \quad \text{for each } R \in \mathcal{AR}.$$

Proposition 4. *For $R \in \mathcal{AR}$,*

$$\#(R) \ = \ \begin{cases} 41 & \text{if } R \text{ is long} \\ 25 & \text{if } R \text{ is medium} \\ 13 & \text{if } R \text{ is e.} \end{cases}$$

Proposition 4 characterizes short, medium and long Allen relations in terms of a notion $\#(R)$ that does not explicitly mention interval borders. The same sum $\#(R)$ arises down the column of the transitivity table

$$\#(R) \ = \ \sum_{R' \in \mathcal{AR}} \text{cardinality}(t(R', R))$$

and is the cardinality of the set

$$\{(R', R'') \in \mathcal{AR} \times \mathcal{AR} \mid R'' \in t(R, R')\}.$$

In the next section, t-numbers $\#(R)$ are built into the probabilites assigned to Allen relations R when three or more intervals are considered.

3 PROBABILITIES OVER n INTERVAL NAMES

The sample space Ω_n in section 2 fixes the number n of linearly ordered points. An alternative is to let $n \geq 2$ be the number of intervals under consideration, construing each element i of $[n]$ not as a point but as an interval. Following (Allen, 1983), we might redefine an n-world to be a function

$$\omega : ([n] \times [n]) \to \mathcal{AR}$$

that labels every pair (i, j) from $[n] \times [n]$ with an Allen relation $\omega(i, j) \in \mathcal{AR}$ in a consistent manner.[2] Consistency of ω here can be understood as the existence of functions

$$\alpha : [n] \to [2n]$$

and

$$\beta : [n] \to [2n]$$

such that for all $i \in [n]$,

$$\alpha(i) < \beta(i) \tag{1}$$

and for all $j \in [n]$,

$$\omega(i, j) \text{ is the Allen relation } R \text{ such that } (\alpha(i), \beta(i)]\ R\ (\alpha(j), \beta(j)]. \tag{2}$$

Together, (1) and (2) turn $i, j \in [n]$ into $<_{2n}$-intervals $(\alpha(i), \beta(i)]$ and $(\alpha(j), \beta(j)]$ that satisfy the specification encoded by ω. The functions α and β above need not be unique, as $[2n]$ may offer plenty of room to satisfy (1) and (2). An extreme example is where all intervals in $[n]$ are equal

$$\omega(i, j) = \ \mathrm{e} \quad \text{for all } i, j \in [n] \tag{3}$$

in which case there are $\binom{2n}{2}$ pairs

$$\alpha, \beta : [n] \to [2n]$$

that work. At the other extreme, exactly one such pair satisfies ω if each interval $i < n$ is before $i + 1$

$$\omega(i, i+1) = \ \mathrm{b} \quad \text{for } i \in [n-1]. \tag{4}$$

These two extreme examples make clear that n is the number of interval names, as opposed to intervals. In the former case, (3), there is just one interval; in the latter, (4), there are un-named intervals between those named in $[n]$. Should we not insist that n count intervals and not just some names? But what, in the finite case, are intervals other than pairs of endpoints? Counting these pairs would lead us back to section 2, with $\binom{k}{2}$ many intervals from k points (give or take 1, for bounds explained in the Introduction). Moreover, it bears noting that interval names are events, which are important ingredients in not only philosophical reconstructions of time but also natural language semantics (Kamp and Reyle, 1993; Kamp, 2013).

For a handle on consistent labellings $\omega : [n] \times [n] \to \mathcal{AR}$, we turn to strings of sets. Recall from Table 1, the strings $\mathfrak{s}_R$ for Allen relations R, such as the string

$$\mathfrak{s}_{\mathrm{m}} = \boxed{l}\boxed{r, l'}\boxed{r'}$$

of length 3, the middle symbol of which is the set with r and l' as its elements. It will be crucial below not to conflate the notions l, l', r, r' even when, as with r and l' in the middle box of $\mathfrak{s}_{\mathrm{m}}$, they name the same point.

Reconstrual of l, l', r, r' in Table 1. *The letters l, l', r and r' appearing in the strings $\mathfrak{s}_R$ in Table 1 are uninterpreted terms (e.g., variables), each distinct from the other (whether or not they co-occur in a box of a string).*

We draw boxes instead of curly braces $\{\cdot\}$ so as not to confuse string symbols with sets such as

$$\{\mathfrak{s}_R \mid R \in \mathcal{AR}\}$$

[2] Subsets of $\mathcal{AR}$ assigned to edges between intervals in (Allen, 1983) are reduced to singletons to keep worlds disjoint, and avoid double counting when basing probabilities on world counts.

which we can form from $\boxed{l}\boxed{r}$ and $\boxed{l'}\boxed{r'}$ through a certain ternary relation & on strings s of sets

$$\&(\boxed{l}\boxed{r},\ \boxed{l'}\boxed{r'},\ s) \iff s \in \{\mathfrak{s}_R \mid R \in \mathcal{AR}\}. \quad (5)$$

(5) is a consequence of defining & by induction according to

$$\text{(i0)}\ \frac{}{\&(\varepsilon,\varepsilon,\varepsilon)} \qquad \text{(i1)}\ \frac{\&(s,s',s'')}{\&(sa,s'a',s''(a\cup a'))}$$

$$\text{(i2)}\ \frac{\&(s,s',s'')}{\&(sa,s',s''a)} \qquad \text{(i3)}\ \frac{\&(s,s',s'')}{\&(s,s'a',s''a')}$$

where ε is the empty string, and a,a' are sets, qua string symbols (Fernando, 2018).[3] The base case (i0) puts $(\varepsilon,\varepsilon,\varepsilon)$ into &, which is closed under rules (i1) for superposition, and (i2), (i3) for shuffling. For example,

$$\&(\boxed{l}\boxed{r},\ \boxed{l'}\boxed{r'},\ \mathfrak{s}_{\mathrm{m}})$$

follows from (i0), (i2), (i1) and (i3)

$$\overset{\text{(i0)}}{\rightsquigarrow}\ (\varepsilon,\varepsilon,\varepsilon)\ \overset{\text{(i2)}}{\rightsquigarrow}\ (\boxed{l},\ \varepsilon, \boxed{l})\ \overset{\text{(i1)}}{\rightsquigarrow}\ (\boxed{l}\boxed{r},\ \boxed{l'},\ \boxed{l}\boxed{r,l'})$$

$$\overset{\text{(i3)}}{\rightsquigarrow}\ (\boxed{l}\boxed{r},\ \boxed{l'}\boxed{r'},\ \boxed{l}\boxed{r,l'}\boxed{r'}).$$

Collecting strings into sets (i.e., languages), we can express & as a binary operation on languages L,L', defining

$$L\&L' := \{s'' \mid (\exists s \in L)(\exists s' \in L')\ \&(s,s',s'')\}\ .$$

We apply & repeateadly to form languages $\mathcal{L}_n$ encoding consistent labellings $\omega : [n]\times[n] \to \mathcal{AR}$. Let

$$\mathcal{L}_1 := \boxed{1}\boxed{1}$$

(following the custom of conflating a string s with the singleton language $\{s\}$) and

$$\mathcal{L}_{n+1} := \mathcal{L}_n \ \&\ \boxed{n+1}\boxed{n+1} \qquad \text{for } n \geq 1.$$

To see how $\mathcal{L}_n$ encodes consistent labellings, a few definitions are in order. Given a set X and a string $s = a_1 \cdots a_k$ of sets,

(i) the X*-reduct* $\rho_X(s)$ *of* s is its componentwise intersection with X

$$\rho_X(a_1\cdots a_k) := (a_1 \cap X)\cdots(a_k \cap X)$$

(ii) the X*-projection* $\pi_X(s)$ *of* s is the result of deleting all occurrences of the empty box $\square$ in $\rho_X(s)$

(Durand and Schwer, 2008). For example,

$$\rho_{\{2,3\}}(\boxed{1,2,4}\boxed{1}\boxed{2,3}\boxed{3}\boxed{4}) = \boxed{2}\boxed{}\boxed{2,3}\boxed{3}\boxed{}$$

$$\pi_{\{2,3\}}(\boxed{1,2,4}\boxed{1}\boxed{2,3}\boxed{3}\boxed{4}) = \boxed{2}\boxed{2,3}\boxed{3}$$

and for any string s, 3 occurs exactly twice in s if $\pi_{\{3\}}(s) = \boxed{3}\boxed{3}$. Clearly, $\mathcal{L}_n$ is the set

$$\{s \in (2^{[n]} - \{\square\})^+ \mid (\forall i \in [n])\ \pi_{\{i\}}(s) = \boxed{i}\boxed{i}\}$$

of strings of non-empty subsets of $[n]$ where each $i \in [n]$ occurs exactly twice. Next, for distinct $i,j \in [n]$ and $R \in \mathcal{AR}$, we let $\mathfrak{s}_{R/i,j}$ be the string $\mathfrak{s}_R$ (from Table 1) with l,r replaced by i, and l',r' replaced by j. For example,

$$\mathfrak{s}_{\mathrm{m}/2,3} = \boxed{2}\boxed{2,3}\boxed{3} \quad \text{and} \quad \mathfrak{s}_{\mathrm{e}/1,2} = \boxed{1,2}\boxed{1,2}$$

and

$$\mathcal{L}_2 = \{\mathfrak{s}_{R/1,2} \mid R \in \mathcal{AR}\}.$$

For $i \neq j$, we can always invert $\mathfrak{s}_R \mapsto \mathfrak{s}_{R/i,j}$ because i and j each occur exactly twice in $\mathfrak{s}_{R/i,j}$. If $s = a_1 \cdots a_k \in \mathcal{L}_n$, and $i,j \in [n]$, then

$$\pi_{\{i,j\}}(s) = \mathfrak{s}_{R/i,j} \iff (l,r]\ R\ (l',r']$$

where l,r are positions in s marked by i

$$l := (\text{least } p \in [k])\ i \in a_p$$
$$r := (\text{greatest } p \in [k])\ i \in a_p$$

and similarly for l',r' and j

$$l' := (\text{least } p \in [k])\ j \in a_p$$
$$r' := (\text{greatest } p \in [k])\ j \in a_p.$$

Accordingly, let us agree s *satisfies* iRj if its $\{i,j\}$-projection is $\mathfrak{s}_{R/i,j}$

$$s \models iRj \iff \pi_{\{i,j\}}(s) = \mathfrak{s}_{R/i,j}.$$

Proposition 5. *Let* $n \geq 2$.

(i) *For all* $s \in \mathcal{L}_n$ *and* $(i,j) \in [n]\times[n]$, *there is a unique* $R \in \mathcal{AR}$ *such that* $s \models iRj$.

(ii) *For all* $s \in \mathcal{L}_n$, *let* $\omega_s : [n]\times[n] \to \mathcal{AR}$ *be the function that sends* (i,j) *to the unique* $R \in \mathcal{AR}$ *such that* $s \models iRj$ *(given by part (i)). The map* $s \mapsto \omega_s$ *is a bijection from* $\mathcal{L}_n$ *onto the set of consistent labellings from* $[n]\times[n]$ *to* $\mathcal{AR}$.

Proposition 5 follows by induction on n. Henceforth, we adopt $\mathcal{L}_n$ as our official sample space, equating the probability of R (for each $R \in \mathcal{AR}$) with the proportion of $\mathcal{L}_n$ in which interval 1 is R-related to interval 2

$$p_n(R) := \frac{\text{cardinality}(\mathcal{L}_n(R))}{\text{cardinality}(\mathcal{L}_n)} \qquad (6)$$

[3] A special case, mix, of the join operation in (Durand and Schwer, 2008) suffices for an unmarked version of (5). The calculation of probabilities below is, however, based on (i0)–(i3).

where $\mathcal{L}_n(R)$ is the subset

$$\mathcal{L}_n(R) := \{s \in \mathcal{L}_n \mid s \models 1R2\}$$

of $\mathcal{L}_n$ satisfying R. The languages $\mathcal{L}_n(R)$ vary with $R \in \mathcal{AR}$, but have a common part (in a sense to be made precise presently), the language $\mathcal{L}_{3:n}$, defined as follows

$$\begin{aligned} \mathcal{L}_{3:2} &:= \varepsilon \\ \mathcal{L}_{3:n+1} &:= \mathcal{L}_{3:n} \,\&\, \boxed{n+1}\boxed{n+1} \qquad \text{for } n \geq 2. \end{aligned}$$

Note that ε is the identity of the binary operation &, which is associative and commutative.

Proposition 6. *For $n \geq 2$, and $R \in \mathcal{AR}$,*

$$\mathcal{L}_n(R) \;=\; \mathfrak{s}_{R/1,2} \,\&\, \mathcal{L}_{3:n}.$$

Behind Proposition 6 is a relationship between & and π_X that can be explained with a couple more definitions. An *X-component of* a string s of sets is a string s' of subsets of X such that

$$\&(s', s'', s) \quad \text{for some string } s'' \text{ of subsets disjoint from } X.$$

We say s is *an S-word* (Durand and Schwer, 2008) if $\square$ does not occur as a symbol in s — i.e.,

$$s \;=\; \pi_{voc(s)}(s)$$

where the *vocabulary $voc(s)$ of s* is the least set X such that $s \in (2^X)^*$

$$voc(a_1 \cdots a_n) \;=\; \bigcup_{i=1}^{n} a_i.$$

Lemma 7. *For all strings s of sets, and disjoint sets X and Y,*

$$\&(\pi_X(s), \pi_Y(s), \pi_{X\cup Y}(s)) \quad (\textit{when } X \cap Y = \emptyset)$$

and if s is an S-word, then $\pi_X(s)$ is the unique S-word that is an X-component of s.

X-components of S-words need not be S-words (e.g., $\boxed{1}\boxed{}$ is a $\{1\}$-component of $\boxed{1}\boxed{2}$) but they are unique after deleting $\square$.

Proposition 8. *Let $n \geq 2$ and s be a string of length $k > 1$ with $n \notin voc(s)$. The set*

$$s \,\&\, \boxed{n}\boxed{n}$$

consists of strings of length k, $k+1$, and $k+2$, of which there are exactly

$$\begin{aligned} d_0(k) &:= \frac{k(k-1)}{2} \;\; \textit{strings of length } k, \\ d_1(k) &:= k(k+1) \;\; \textit{strings of length } k+1, \textit{ and} \\ d_2(k) &:= \frac{(k+1)(k+2)}{2} \;\; \textit{strings of length } k+2. \end{aligned}$$

A string in $s\&\boxed{n}\boxed{n}$ of length k chooses 2 positions from s in which to put n, whence

$$d_0(k) = \binom{k}{2}$$

while length $k+1$ chooses a position from s and one of $k+1$ positions not in s

$$d_1(k) = k(k+1)$$

and length $k+2$ chooses 2 positions outside s, which may be different or the same

$$d_2(k) = \binom{k+1}{2} + k + 1 = \frac{(k+1)(k+2)}{2}.$$

Returning now to the probabilities defined by line (6) above, let $c_n(R)$ be the number

$$c_n(R) \;:=\; \text{cardinality}(\mathcal{L}_n(R))$$

of strings in $\mathcal{L}_n$ satisfying R. It is instructive to observe that $c_3(R)$ is just the t-number $\#(R)$ defined at the end of section 2 as the sum of the transitivity table row for R

$$c_3(R) \;=\; \sum_{R' \in \mathcal{AR}} \text{cardinality}(t(R, R')).$$

For all $n \geq 2$, we can calculate the quantities $c_n(R)$ in terms of

$$c_n(R;k) \;:=\; \text{cardinality}(\{s \in \mathcal{L}_n(R) \mid \text{length}(s) = k\})$$

for which we have the recurrence

$$c_2(R;k) \;=\; \begin{cases} 1 & \text{if length}(\mathfrak{s}_R) = k \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$\begin{aligned} c_{n+1}(R;k) &= c_n(R;k)d_0(k) + c_n(R;k-1)d_1(k-1) \\ &\quad + c_n(R;k-2)d_2(k-2) \\ &= d_0(k)(c_n(R;k) + 2c_n(R;k-1) \\ &\quad + c_n(R;k-2)) \end{aligned} \tag{8}$$

from Proposition 8, with Lemma 7 ruling out the possibility that (8) double counts. Propositions 6 and 8 reduce the variation in $p_n(R)$ to the length of $\mathfrak{s}_R$

$$c_n(R) = c_n(R') \;\text{ if }\; \text{length}(\mathfrak{s}_R) = \text{length}(\mathfrak{s}_{R'})$$

for all $R, R' \in \mathcal{AR}$ and $n \geq 2$. For the record,

Table 3: Some probabilities of e, m, b.

n	$p_n(\text{e})$	$p_n(\text{m})$	$p_n(\text{b})$	γ_n	γ'_n	$1-6p_n(\text{b})$
2	$\frac{1}{13}$	$\frac{1}{13}$	$\frac{1}{13}$	1	1	$\frac{7}{13} \approx 0.538461538$
3	0.031784841	0.061124694	0.100244499	2	2	0.398533007
10	0.002527761	0.021841026	0.144404347	9	7	0.133573915
100	0.000023782	0.002283051	0.164379652	96	72	0.013722086
500	0.000000959	0.000460405	0.166206102	480	361	0.002763387
1000	0.000000240	0.000230840	0.166435786	961	721	0.001385281
1500	0.000000107	0.000153893	0.166512755	1442	1082	0.000923468

Theorem 9. *For $n \geq 2$ and $R \in \mathcal{AR}$, the probabilities $p_n(R) = c_n(R)/c_n$ can be calculated as follows*

$$c_n(\text{e}) = \sum_{k=2}^{2n-2} c_n(\text{e};k) \tag{9}$$

$$c_n(R) = \sum_{k=3}^{2n-1} c_n(R;k) \quad \textit{for medium } R$$

$$c_n(R) = \sum_{k=4}^{2n} c_n(R;k) \quad \textit{for long } R$$

where $c_n(R;k)$ is given by lines (7) and (8) above, and

$$c_n = c_n(\text{e}) + 6(c_n(\text{m}) + c_n(\text{b})) \tag{10}$$

(representing medium relations by meet, m*, and long relations by before,* b*).*

The summation index k in Theorem 9 ranges over the possible lengths of strings in $\mathcal{L}_n(R)$, according to whether R is short, medium or long. One can map the language $\mathcal{L}_n$ to $\mathcal{L}_{n+1}(\text{e})$ by a bijection that renames interval i to $i+1$ and inserts 1e2, establishing

$$c_n = c_{n+1}(\text{e}).$$

Hence, as an alternative to (9), we can specify $c_n(e)$ by the recurrence

$$c_2(\text{e}) = 1 \quad (= c_2(\text{m}) = c_2(\text{b}))$$
$$c_{n+1}(\text{e}) = c_n(\text{e}) + 6(c_n(\text{m}) + c_n(\text{b})) \quad \text{for } n \geq 2.$$

It is (9) and $c_n(\text{e};k)$, however, that appear in Sloane's On-line Encyclopedia of Integer Sequences for the "number of different relations between n intervals on a line"

$$a(n) = \sum_{i=2}^{2n} \lambda(i,n) \quad \text{where } \lambda(i,n) = c_n(\text{e};i)$$
$$\text{(according to (7), (8) above)}$$

in `https://oeis.org/A055203`.[4]

[4] It is conjectured there that $a(n) = 1 \bmod 12$, which is equivalent to the claim that $c_n(\text{m}) + c_n(\text{b})$ is even, by (10) in Theorem 9.

Some values of $p_n(R)$ are listed in Table 3, alongside integers γ_n and γ'_n that compare $p_n(\text{m})$ to $p_n(\text{e})$

$$\gamma_n := \left\lceil \frac{p_n(\text{m})}{p_n(\text{e})} \right\rceil = \left\lceil \frac{c_n(\text{m})}{c_n(\text{e})} \right\rceil$$

and $p_n(\text{b})$ to $p_n(\text{m})$,

$$\gamma'_n := \left\lceil \frac{p_n(\text{b})}{p_n(\text{m})} \right\rceil = \left\lceil \frac{c_n(\text{b})}{c_n(\text{m})} \right\rceil$$

respectively. The inequalities

$$\frac{c_n(\text{m})}{c_n(\text{e})} < \frac{c_{n+1}(\text{m})}{c_{n+1}(\text{e})}$$

and

$$\frac{c_n(\text{b})}{c_n(\text{m})} < \frac{c_{n+1}(\text{b})}{c_{n+1}(\text{m})}$$

have been verified computationally for $2 \leq n \leq 1500$, providing evidence but *not* a proof that the asymptotic probabilities described in Corollary 3 carry over to $\mathcal{L}_n$. The case $n = 2$ reproduces our first answer to the question (Q) in the Introduction above

$$p_2(R) = \frac{1}{13}$$

while the transitivity table numbers $\#(R)$ are the basis for $n = 3$

$$p_3(R) = \frac{\#(R)}{\sum_{R' \in \mathcal{AR}} \#(R')}$$

which varies according to whether R is short, medium or long.

4 DISCUSSION

The study of probabilities above has led us to partition Allen relations between the short, medium and long, which is far less common than that between overlap

$$\bigcirc = \bigvee\{\text{d,di,o,oi,s,si,f,fi,e}\},$$

precedence

$$\prec = \bigvee\{\text{m,b}\},$$

and its converse

$$\succ \; = \; \bigvee\{\text{mi,bi}\}$$

(Kamp and Reyle, 1993; Durand and Schwer, 2008, among others). Using section 2, the asymptotic probabilities

$$\begin{aligned} p(\bigcirc) &= \lim_{n\to\infty} p_n(\text{d}) + p_n(\text{di}) + p_n(\text{o}) + p_n(\text{oi}) + \\ &\qquad p_n(\text{s}) + p_n(\text{si}) + p_n(\text{f}) + p_n(\text{fi}) + p_n(\text{e}) \\ &= \frac{2}{3} \\ p(\prec) &= \lim_{n\to\infty} p_n(\text{m}) + p_n(\text{b}) = \frac{1}{6} \\ p(\succ) &= \lim_{n\to\infty} p_n(\text{mi}) + p_n(\text{bi}) = \frac{1}{6} \end{aligned}$$

do not differ vastly from the numbers

$$9/13, \;\; 2/13, \;\; 2/13$$

obtained by replacing the probabilities $p_n(R)$ of an Allen relation R uniformly with $1/13$, the probability $p_2(R)$ where section 3 starts (at $n = 2$). While variations in n are of limited consequence for $\bigcirc$, $\prec$ and $\succ$, it is a another matter once $\bigcirc$, $\prec$ and $\succ$ are refined to Allen relations. But why invite such complications?

An important reason to be interested in n is granularity, which takes on particular significance when it is varied. One way to see this is through Leibniz's law, indiscernibility as identity. The requirement that any difference $x \neq y$ is discernible via some property P can be expressed in monadic second-order logic (Libkin, 2010, for example) as

$$x \neq y \;\supset\; (\exists P)\neg(P(x) \equiv P(y)). \tag{LL}$$

If we replace $\neq$ by adjacency S and restrict P to be given by some finite set X, (LL) becomes "time steps$_S$ require change$_X$"

$$xSy \;\supset\; x \not\equiv_X y \tag{LL$_{S,X}$}$$

where $x \not\equiv_X y$ means: x and y differ over some predicate from X

$$x \not\equiv_X y \;:=\; \bigvee_{i\in X} \neg(P_i(x) \equiv P_i(y)).$$

For each $i \in X$, let us mark P_i's left and right borders with subscripts $l(i)$ and $r(i)$ for predicates $P_{l(i)}$ saying: P_i is false but S-after true

$$P_{l(i)}(x) \;\equiv\; \neg P_i(x) \wedge (\exists y)(xSy \wedge P_i(y)) \tag{11}$$

and $P_{r(i)}$ saying: P_i is true but not S-after

$$P_{r(i)}(x) \;\equiv\; P_i(x) \wedge \neg(\exists y)(xSy \wedge P_i(y)). \tag{12}$$

Formulating $x \not\equiv_X y$ as

$$\bigvee_{i\in X}((\neg P_i(x) \wedge P_i(y)) \vee (P_i(x) \wedge \neg P_i(y))$$

brings us, under xSy, to $\bigvee_{i\in X}(P_{l(i)}(x) \vee P_{r(i)}(x)))$

$$xSy \;\supset\; (x \not\equiv_X y \;\equiv\; \bigvee_{i\in X}(P_{l(i)}(x) \vee P_{r(i)}(x)))$$

assuming (11), (12) and S is deterministic

$$(\forall z)(xSy \wedge xSz \supset y = z). \tag{13}$$

That is, under (11)–(13), (LL$_{S,X}$) says:

$$(\exists y)(xSy) \;\supset\; \bigvee_{i\in X}(P_{l(i)}(x) \vee P_{r(i)}(x)). \tag{14}$$

To enforce (14), we let $X_\bullet$ be the set

$$X_\bullet \;:=\; \{l(i) \mid i \in X\} \cup \{r(i) \mid i \in X\}$$

of borders in X, and define a translation

$$\beta : (2^X)^* \to (2^{X_\bullet})^*$$

with for example,

$$\beta(\boxed{}\boxed{i,i'}\boxed{i'}) = \boxed{l(i),l(i')}\boxed{r(i)}\boxed{r(i')}$$

mapping, in general, a string $a_1 \cdots a_k$ of subsets of X to the string $b_1 \cdots b_k$ of subsets of $X_\bullet$ according to (11) and (12)

$$\begin{aligned} b_x &:= \{l(i) \mid i \in a_{x+1} - a_x\} \cup \\ &\quad\;\; \{r(i) \mid i \in a_x - a_{x+1}\} \qquad \text{for } x < k \\ b_k &:= \{r(i) \mid i \in a_k\} \end{aligned} \tag{15}$$

(Fernando, 2018). While (13) is built into every string, (14) is not. For a non-final position x, (15) says

$$\begin{aligned} b_x \neq \square &\iff (a_{x+1} - a_x) \cup (a_x - a_{x+1}) \neq \square \\ &\iff a_{x+1} \neq a_x. \end{aligned}$$

That is, for $b_1 \cdots b_k = \beta(a_1 \cdots a_k)$,

$b_1 \cdots b_{k-1}$ is an S-word $\iff$ $a_1 \cdots a_k$ has *no* stutter

where a *stutter* of $a_1 \cdots a_k$ is a non-final position $x \in [k-1]$ such that

$$a_x = a_{x+1}.$$

An S-word $\beta(s)$ satisfies (14) and a bit more

$$(\forall x)\bigvee_{i\in X}(P_{l(i)}(x) \vee P_{r(i)}(x))$$

without the precondition

$$(\exists y)(xSy)$$

that x is not S-final.

For each Allen relation R, we can picture $1R2$ not only as the S-word $\mathfrak{s}_{R/1,2}$ from Table 1, but also as a stutterless string $\mathfrak{s}_R^\circ$ in Table 4 (Fernando, 2016, page 3635), stepping outside S-words for

$$\mathfrak{s}_{\text{b}}^\circ = \boxed{1}\boxed{}\boxed{2}$$

and

$$\mathfrak{s}_{\text{bi}}^\circ = \boxed{2}\boxed{}\boxed{1}.$$

Table 4: Allen relations via stutterless strings.

R	$\mathfrak{s}_R^\circ$	R^{-1}	$\mathfrak{s}_{R^{-1}}^\circ$
b	$\boxed{1}\boxed{}\boxed{2}$	bi	$\boxed{2}\boxed{}\boxed{1}$
o	$\boxed{1}\boxed{1,2}\boxed{2}$	oi	$\boxed{2}\boxed{1,2}\boxed{1}$
m	$\boxed{1}\boxed{2}$	mi	$\boxed{2}\boxed{1}$
d	$\boxed{2}\boxed{1,2}\boxed{2}$	di	$\boxed{1}\boxed{1,2}\boxed{1}$
s	$\boxed{1,2}\boxed{2}$	si	$\boxed{1,2}\boxed{1}$
f	$\boxed{2}\boxed{1,2}$	fi	$\boxed{1}\boxed{1,2}$
e	$\boxed{1,2}$		

From Table 4, Table 1 is a small step away

$$\mathfrak{s}_R \approx \beta(\Box\mathfrak{s}_R^\circ) \text{ for } l \approx l(1),\ r \approx r(1),$$
$$l' \approx l(2),\ r' \approx r(2).$$

For example, $R = \text{m}$ gives

$$\beta(\boxed{}\boxed{1}\boxed{2}) = \boxed{l(1)}\boxed{r(1),l(2)}\boxed{r(2)}$$
$$\approx \boxed{l}\boxed{r,l'}\boxed{r'}.$$

Stutterless strings arise from de-stuttering

$$saas' \rightsquigarrow sas' \tag{16}$$

just as S-words arise from $\Box$-removal

$$s\Box s' \rightsquigarrow ss'. \tag{17}$$

(17) implements the Aristotelian slogan

no time without change

under the assumption that

(†) all predicates in a string symbol a express change.

By contrast, (16) reflects the assumption that strings are built from cumulative predicates, where by definition, a predicate P on intervals is *cumulative* if whenever an interval i meets an interval i' for the combined interval $i \sqcup i'$,

$$P(i) \text{ and } P(i') \implies P(i \sqcup i').$$

The converse

$$P(i \sqcup i') \implies P(i) \text{ and } P(i')$$

(for i meets i') is what it means for P to be *divisive*. P is cumulative and divisive precisely if it satisfies the condition (H) for *homogeneity*

(H) for all intervals i and i' whose union $i \cup i'$ is an interval,

$$P(i \cup i') \iff P(i) \text{ and } P(i').$$

A bias towards stutterless strings (as opposed to S-words) is in line with the well-known aspect hypothesis from (Dowty, 1979) claiming

> the different aspectual properties of the various kinds of verbs can be explained by postulating a single homogeneous class of predicates — stative predicates — plus three or four sentential operators or connectives. (page 71)

That said, it is no accident that non-stative borders are strung together in Table 1 for use in both sections 2 and 3, whereas their stative interiors are relegated (for present purposes) to Table 4. Our analysis of Allen relations above focuses not on the static condition of interiors (described by (H)), but on the change marked by borders (in accordance with (†)).

There are reasons to shift the aforementioned focus towards a more even balance in future work. Statives and non-statives are boxed together in *discourse representation structures* (Kamp and Reyle, 1993), which can be put one after another in strings to describe regularities (such as the preconditions and effects of actions) beyond chance. Chance is assessed above relative to sample spaces Ω_n consisting of worlds linked to model-theoretic interpretations of discourse representation structures. These model-theoretic interpretations can be recast in ordinary predicate logic, on which probabilities can be defined. An equation assigning probabilities $p(x)$ to worlds x that has received considerable attention in recent years is

$$p(x) = \frac{1}{Z}\, exp(\sum_{i \in I} w_i n_i(x)) \tag{18}$$

(Domingos and Lowd, 2009) given some finite set I of first-order formulas i and weights $w_i \in \mathbb{R}$ that shape the probability of x according to the number $n_i(x)$ of groundings in x that satisfy i. (18) is applied to interval networks for event recognition in (Morariu and Davis, 2011), one of a number of works with data-driven assignments of probabilities to Allen relations (Zhang et al., 2013; Liu et al., 2018, among others). The contribution of I to (18) is neutralized if every weight w_i is 0 (or equivalently, $I = \emptyset$), resulting in equiprobable worlds (with Z in (18) equal to the number of such worlds). It is this null, data-free case on which we focus when raising in the Introduction the question (Q) of the probability of aRa', for *arbitrary* intervals a, a'. Our answers, Theorem 2 in section 2 and Theorem 9 in section 3, are based on finite sample spaces Ω_n of temporal entities that divide Allen relations into the short, medium and long. No previous attention has, as far as we know, been paid to this division. Does the division fade into insignificance once an account of actions is introduced through a non-empty set I of formulas and non-zero weights in (18)? That would depend on I, which we have put aside in answering (Q).

5 CONCLUSION

The probability an Allen relation holds between two arbitrary intervals is specified in Theorems 2 and 9 under the assumption that intervals are drawn from a finite model by a fair method (in accordance with the principle of indifference). The finite model assumed depends on the particular application at hand. (For example, the passage above from (Kamp, 2013) describes a range of applications where that model is based on the events mentioned in a discourse.) Whether or not the notion of a fair coin can or should extend to the choice of intervals from any such model is a natural question that, in our view, merits study.

ACKNOWLEDGEMENTS

We are grateful to three anonymous referees for their helpful comments.

This research is supported by Science Foundation Ireland (SFI) through the CNGL Programme (Grant 12/CE/I2267) in the ADAPT Centre (https://www.adaptcentre.ie) at Trinity College Dublin. The ADAPT Centre for Digital Content Technology is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

REFERENCES

Allen, J. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843.

Allen, J. and Ferguson, G. (1994). Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4(5):531–579.

Domingos, P. and Lowd, D. (2009). *Markov Logic: An Interface Layer for Artificial Intelligence*. Morgan and Claypool Publishers.

Dowty, D. (1979). *Word Meaning and Montague Grammar*. Reidel, Dordrecht.

Durand, I. and Schwer, S. (2008). A tool for reasoning about qualitative temporal information: the theory of S-languages with a Lisp implementation. *Journal of Universal Computer Science*, 14(20):3282–3306.

Fernando, T. (2016). Prior and temporal sequences for natural language. *Synthese*, 193(11):3625–3637.

Fernando, T. (2018). Intervals and events with and without points. In *Proceedings of the Symposium on Logic and Algorithms in Computational Linguistics 2018*, pages 34–46. Stockholm University DiVA Portal for digital publications.

Freksa, C. (1992). Temporal reasoning based on semi-intervals. *Artificial Intelligence*, 54:199–227.

Hamblin, C. (1971). Instants and intervals. *Studium generale*, 24:127–134.

Kamp, H. (2013). The time of my life. https://lucian.uchicago.edu/blogs/elucidations/files/2013/08/Kamp_TheTimeOfMyLife.pdf.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic*. Kluwer.

Libkin, L. (2010). *Elements of Finite Model Theory*. Springer.

Liu, L., Wang, S., Hu, B., Qiong, Q., Wen, J., and Rosenblum, D. (2018). Learning structures of interval-based Bayesian networks in probabilistic generative model for human complex activity recognition. *Pattern Recognition*, 81:545–561.

Morariu, V. and Davis, L. (2011). Multi-agent event recognition in structured scenarios. *Proc.IEEE Conf. Computer Vision and Pattern Recognition*, pages 3289–3296.

Schwer, S. (Last modified Dec 2018). Sequence A055203 in *The On-Line Encyclopedia of Integer Sequences*. https://oeis.org/A055203.

Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Moszkowicz, J., and Pustejovsky, J. (2009). The TempEval Challenge: Identifying temporal relations in text. *Language Resources and Evaluation*, 43(2):161–179.

Zhang, Y., Zhang, Y., Swears, E., Larios, N., Wang, Z., and Ji, Q. (2013). Modeling temporal interactions with interval temporal Bayesian networks for complex activity recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(10):24682483.