

Temporal Vagueness, Coordination and Communication

Ewan Klein and Michael Rovatsos

School of Informatics, University of Edinburgh

1 Introduction

How is it that people manage to communicate even when they implicitly differ on the meaning of the terms they use? Take an innocent-sounding expression such as *tomorrow morning*. What counts as *morning*? There is a surprising amount of variation across different people.¹

For Anna, morning starts ‘when she gets up’, and finishes ‘when she has lunch’. For Bart (who verges on the pedantic), morning officially starts at 12:00 am and ends at 11:59 am. Yet another view, held by Cecile, is that morning starts sometime between 6:00 and 7:00 am, and ends sometime between 12:30 pm and 1:30 pm. Finally, Devendra (who regularly works into the small hours) believes that morning has barely started at 10:30 am and finishes around 3:30 pm. Nevertheless, if Anna says to Bart: *drop by my office tomorrow morning and we’ll have a look at your proposal*, the chances are high that Anna and Bart will manage to meet (as long as they have no conflicting engagements).

In the kind of linguistic contexts we are concerned with in this paper, it seems plausible to treat *morning* as a grouping of time units at some level of granularity (e.g., seconds, minutes, quarter-hours), ordered in the usual way. According to this view, a sentence like *Let’s meet tomorrow morning* is equivalent to *Let’s meet at some point in tomorrow morning*. This allows us to claim, for example, that the moment 9:15 am belongs to the extension of *morning*, while 9:15 pm does not. It follows that *morning* is open to the Sorites paradox: if 9:15:00 am counts as morning, then so does 9:15:01 am (i.e., the moment that is one second later than 9:15:00 am). By tediously iterating through the process of adding one second at a time (or one millisecond, if preferred), we will ineluctably reach the unwanted conclusion that 9:15:00 pm counts as morning. If we take the Sorites paradox as criterial for vagueness, we can conclude that *morning* and its companion expressions, *afternoon*, *evening*, *day* and *night* are all vague. But what does this mean? On the face of it, some speakers (like Bart) assign crisp boundaries to the time unit *morning*, while others (like Cecile) assign indeterminate boundaries. We will return to this issue in Section 3.3, but for the time being, let us just assume that the concepts corresponding to familiar time units possess crisp boundaries. Instead, we want to explore how terms like *morning* might be used in communities of speakers.

¹ The variability in usage and interpretation of terms like *morning* and *evening* has been explored by Reiter [18] in the context of weather forecasts.

Vagueness and Utility

The approach we have adopted is inspired in large part by Parikh’s [16] observation that even though two speakers differ in the way they interpret a vague term like *blue*, if there is sufficient overlap in their interpretations, there will be positive utility in using the vague term. In Parikh’s example, Ann requests Bob to fetch “a blue book on topology” from the book shelves in her study. The descriptive term contains enough information that even though they disagree on what counts as blue, the set of ‘blue-for-Bob’ books reduces Bob’s search space far enough to significantly increase his chances of finding the correct book relatively fast.

What we want to adopt from Parikh’s scenario is the idea that the success of communication involving a vague term can be measured in terms of completing a task. In Parikh’s case, the task is to identify a book; in our case, the task is for two agents to meet one another. Just as the term *blue* functions in Parikh’s scenario to reduce the search space within which the required book is located, we will assume that a term like *morning* reduces the temporal period within which the meeting will take place. More specifically, we assume there are two agents, say A_1 and A_2 , who wish to meet up. Suppose A_1 says to A_2 : *Let’s meet up tomorrow morning. Drop by my office.* A_2 accepts the proposal. Both A_1 and A_2 have their own interpretation of what is meant by the phrase *morning*. For each of them, the interpretation is modelled as an interval, but these intervals do not need to coincide. Not surprisingly, we can observe that if the intervals overlap sufficiently, then the two agents will tend to be successful in meeting.

Although we focus in this paper on temporal intervals, in principle we could generalize our approach to any linguistic term whose semantic extension is a set. We define **overlap** between sets as follows:

Definition 1 *The degree of overlap between sets X and Y , $\circ(X, Y)$, is the quotient*

$$\frac{|X \cap Y|}{|X \cup Y|}$$

i.e., the cardinality of elements in the intersection of X and Y divided by the cardinality of elements in the union of X and Y .

If $\circ(X, Y) = 1.0$ then we say that X and Y *completely overlap*. Given some error margin ϵ , we will say that X and Y *approximately overlap* iff $1 - \circ(X, Y) < \epsilon$. In this case, we can say that there is an indifference relation between X and Y : the difference between them is either indiscernable or has no practical impact for the agents.

Let us write $V_i(e)$ for the interpretation that agent A_i assigns to expression e ; we will restrict our attention to cases where $V_i(e)$ is a set $X \subseteq \mathcal{D}$ for some domain \mathcal{D} . Two interpretations V, V' are *completely (resp. approximately) aligned on e* iff $V(e)$ and $V'(e)$ completely (resp. approximately) overlap.

We assume that holding a meeting always has higher utility than failing to meet, that is, $U(\text{meet}) > U(\overline{\text{meet}})$. $P(\text{meet} \mid V_i(e))$ is the probability that a

meeting will take place, given the interpretation that A_i assigns to expression e . More generally, let's assume that S is an event whose occurrence is conditioned by linguistic meanings. Then the *expected utility* of an interpretation V_i , relative to outcome S , is given as:

$$EU(V_i) = P(S \mid V_i(e))U(S) + P(\bar{S} \mid V_i(e))U(\bar{S})$$

If $EU(V'_i) > EU(V_i)$, then a rational agent A_i should adopt V'_i in place of V_i in order to maximize her expected utility. Using these notions, we go beyond Parikh's scenario, and make the following claim:

If individual agents in a given community maximize the expected utility of their interpretations, then over the course of successive interactions these interpretations will become approximately aligned.

Much of the remainder of this paper will attempt to flesh out and substantiate this claim. However, we should emphasize that the model that we develop does not attempt to directly compare the utility of all possible interpretations at a given point in the interaction. Rather, the principle of utility maximization is comparable to an abstract specification which admits various computational implementations.

Conceptual Structures

The standard assumption in formal semantics (and indeed in much computational semantics) holds that linguistic meaning is a mapping from language to the world (or a model): meanings have an objective existence independently of speakers. By making the interpretation function V relative to agents, we are implicitly subscribing to a cognitive view, where meanings are psychological entities in the heads of agents. From the perspective of building some kind of computational system of interacting agents (such as mobile robots), the cognitive approach has obvious attractions. Each agent has only partial knowledge of the world in which it finds itself, including both the physical environment and its fellow agents. It does not have direct access to 'external reality', but has to build representations of the world on the basis of input from its sensors (which may well be noisy). It could be argued that it is enough to equip the agents with a mental language, such as some flavour of first order logic, in order to reason and communicate. But this begs the question of how the agents can be sure that they are using the non-logical terms of the language in the same way as their dialogue partners.

Within Gärdenfors' framework of conceptual spaces [8, 26], concepts (and hence linguistic meanings) are internal mental representations. However, the requirement of 'shareability' [7] places constraints on how far the concepts of one agent can diverge from those of the other agents it interacts with. Shared meanings of expressions develop during language games — communicative interaction between language users — and involve mappings between conceptual representations that are influenced by the need to act effectively in the world. A

so-called ‘meeting of minds’ occurs when the representations in the minds of the dialogue partners become sufficiently compatible. This is essentially the same as our notion of approximate alignment.²

Our approach to meaning is also influenced by work on ontology alignment [5] in the context of multi-agent systems [1]. Agents collaborating in a shared environment need to share an ontology (i.e., the conceptualization of a domain) in order to communicate with each other, but in an open system, different agents can in principle use quite heterogeneous ontologies. Wang and Gasser [24] present a model that, like ours, explicitly considers which instances fall within the extension of a concept, but do not provide a utility-based method for determining successful alignment. Somewhat closer to our approach in this respect is the work of McNeill et al. [14], where agents are involved in jointly planning a task; plan failure triggers an attempt to diagnose mismatches in ontology; the agents use heuristics to repair their ontologies (in the sense of modifying the ontology signature), and then re-engage in the planning task. This cycle — communicate / diagnose failure / repair the ontology — is similar to the kind of model that we are proposing. However, the type of mismatches considered by [14], and the mechanisms used to effect the repair, are very different.

One question which arises is whether it is plausible that agents are prepared to modify their interpretations in the way we have suggested. Although this point deserves closer consideration, it does seem to be a characteristic of vague terms (both adjectives and nouns) that their boundaries are somewhat flexible. Thus, we seem to be more willing to shift the boundaries of what counts as morning than, say, what counts as a dog (or other natural kind).³ On the other hand, even if agents are prepared to ‘negotiate’ meaning, there are no doubt some aspects which are non-negotiable — Bob may be prepared to shift his interpretation of *blue* so that it encompasses a shade of violet, but will balk at shifting it to cover bright orange. This is an important constraint, but we will defer the topic to future work.⁴

Overview of paper

We use a simple multi-agent simulation in order to provide an explicit model of task based communication. In general, we believe this has a number of attractive aspects. For one thing, the simulation allows us to explore the consequences of setting various parameters in different ways, and to consider the interaction of these parameters in a manner that would be hard to achieve using a pencil-and-paper analysis. The approach can be viewed as implementing a language game

² An elegant computational implementation of alignment of colour terms in Gärdenfors’ framework is presented by Jäger and van Rooij [10].

³ This is similar to Williamson’s [25] proposal that the meanings of vague terms are *unstable*, in the sense that minor differences in use give rise to minor differences in the extension of the term.

⁴ For more discussion of constraints on shifting meanings in a computational framework, see [13, 4].

in the sense of Gärdenfors [8], where the representations of individual agents affect communication about shared activities and are modified as a result.

Section 2 describes the framework of the simulation in more detail. Section 3 and Section 4 present the two sets of experiments that we ran, while Section 5 gives some conclusions and suggestions for future work.

2 Approach

As we have already indicated, our treatment of temporal expressions is highly simplified. Most notably, we ignore the element of context dependence in the application of temporal terms. For example, people who work together in an office will probably adopt a different view of what counts as morning than people who are up before dawn to milk the cows. Another contextual factor is the day of the week: for most Westerners, the temporal location of *morning* during the weekend diverges considerably from its location during the working week. We will abstract away from these factors, and only consider the case where the population of speakers adopts a shared context of use.

A second simplification is in our treatment of the expression *morning*. Given a specific day (say Monday 9th November 2009) and a specific speaker, say Anna, *morning* will denote a closed interval of time units.⁵ For our purposes, it does not matter too much what level of granularity is chosen, but we will think of the intervals used by our agents as containing quarter-hour units; in other words, an interval with 12 elements would correspond to a period whose duration is three hours.

We will describe two families of experiments (referred to as Experiment 1 and Experiment 2 respectively), using a multi-agent simulator that was implemented in the Python programming language.⁶ The agents are modelled as processes in the SimPy Discrete Event Simulator.⁷ Before discussing the specifics of the experiments, we will give more details of the agent coordination task.

Let \mathcal{T} be a finite set of integers representing **time units**, and let \mathcal{I} be a set of closed intervals over \mathcal{T} . Given a set Ag of agents, each $A_i \in Ag$ is associated with a **preferred interval** $\iota_i \in \mathcal{I}$. We will assume that $\iota_i = V_i(\textit{morning})$, i.e., A_i 's interpretation of the temporal expression *morning*. $V_i(\textit{morning})$ is private in the sense that for any $j \neq i$, A_j has no direct access to V_i .

Note that although $V_i(\textit{morning})$ is unique for each agent A_i , the inverse need not hold — that is, we let the cardinality of \mathcal{I} be less than that of Ag . In

⁵ This approach is intended to be compatible with that proposed by Ohlbach [15], who points out that a temporal expression such as *February* can be used to refer to a particular February; or to denote the set of all Februaries in the history of mankind; or, more generally, to refer to a function which given some year y returns the particular February of y .

⁶ <http://www.python.org/>

⁷ <http://simpy.sourceforge.net/>

our simulations, \mathcal{I} is fixed as the set of intervals $\{[1, 10], [6, 15], [11, 20]\}$.⁸ It is assumed in our model that the agents share the common time frame given by \mathcal{T} . For example, we might think of the three intervals in \mathcal{I} as corresponding roughly to the time periods 7:00–9:30 am, 8:15–10:30 am and 9:30–12.00 am, respectively, where the time units 7:00 am, 7:15 am, ... have the same interpretation for all agents in Ag .⁹

On each run of a simulation, two agents A_i and A_j are selected at random. One of the agents is assigned the role of **proposer**, while the other takes on the role of **responder**; we'll refer to these as P and R respectively. P takes the lead in sending a "let's meet in the morning" message to R and chooses an arrival time arr_P from its period ι_P , while R chooses an arrival time arr_R from ι_R . One important feature of the model (which could however be relaxed) is that agents tend to pick an arrival time that falls somewhere in the middle of their preferred interval. This seems plausible when the proposed meeting time is some kind of approximation or vague interval. This feature is implemented by selecting A_i 's arrival time (coerced to an integer) at random from a Gaussian distribution whose mean is the midpoint of ι_i , with standard deviation 1. In Experiment 1, the departure time of an agent A_i , dep_i , is simply set to the endpoint of ι_i . (We will later discuss a modification of this scheme used in Experiment 2.) P and R are judged to **meet** if $[arr_P, dep_P] \cap [arr_R, dep_R] \neq \emptyset$. We assume that on each run, P knows the arrival and departure time of R, even if they fail to meet.

It may be helpful to enumerate the four cases which determine whether or not a meeting occurs. (Although we mention a 'waiting cost' here, this feature does not come into play until Experiment 2.)

1. R arrives and departs before arr_P ; the meeting fails with no waiting cost for P.
2. R has already arrived but not yet departed when P arrives; the meeting is accomplished with no waiting cost for P.
3. R arrives after arr_P but before dep_P ; the meeting is accomplished with a waiting cost for P.
4. R arrives after dep_P ; the meeting fails with a waiting cost for P.

These four options are shown graphically in Figure 1.

As mentioned before, each agent is assigned a preferred interval, which is intended to be a cognitive representation of a vague temporal expression. Since there is only one such expression in use in the community, we do not need to explicitly label it. The preferred interval, therefore, is a key aspect of each agent's mental state. Agents have no access to the mental states of others, and only observe their behaviour in arriving and departing at particular times. Each agent keeps a record of the other's arrival behaviour. More precisely, each agent

⁸ These integer bounds are chosen for simplicity of implementation, but it would be conceptually straightforward to replace them with time points in *hh:mm* notation, or indeed to use seconds in Unix time (<http://unixtime.info/>).

⁹ See [19] for discussion of the cultural and cognitive construction of time based time interval systems.

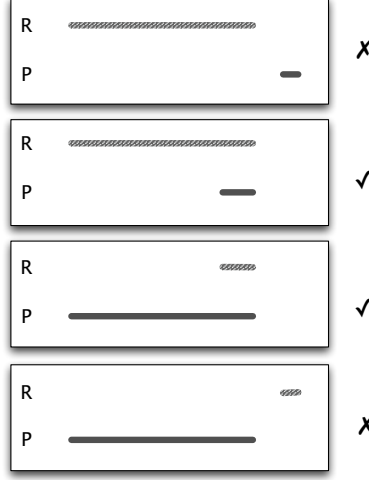


Fig. 1: Meeting Outcomes

A_i maintains a list $L(A_j)$ of observed arrival times for each other agent A_j , and the list is updated on any run in which A_i plays the P role and A_j plays the R role. Given $L(A_j)$, A_i can estimate the mean arrival time of A_j up to the current run in the simulation. We refer to this estimated mean as $\mu(\iota_j)$.

In order to provide a more concrete impression of the way the simulation works, in Figure 2 we have included a small extract from one simulation log file.

3 Experiment 1

3.1 Alignment

In the first set of experiments, we allow the proposer to update its preferred interval in the light of its experience so far. After each encounter, P attempts to **align** with R. It does so by adjusting ι_P so that the midpoint of ι_P approaches $\mu(\iota_R)$; that is, if t is the new target midpoint and len returns the length of an interval, then the adjusted interval is simply $[t - len(\iota_i)/2, t + len(\iota_i)/2]$. Let us refer to the midpoint of interval ι_P as $md(\iota_P)$ and let ι'_P be the new interval of P after alignment has taken place. Then we try to meet the following constraint after each run:

$$|\mu(\iota_R) - md(\iota_P)'| < |\mu(\iota_R) - md(\iota_P)| \quad (1)$$

In Experiment 1, we implemented the following update rule, where $\lambda \in [0, 1]$ is a scaling factor that we call the **learning rate**:

$$md(\iota_P)' = md(\iota_P) + \lambda(\mu(\iota_R) - md(\iota_P)) \quad (2)$$

```

activating agent-2 at 26
agent-2 proposed the following period:
[4, 5, 6, 7, 8, 9, 10, 11, 12, 13]
Proposer: agent-2, Responder: agent-4
Failed to meet!
agent-2 present: [11, 12, 13]
agent-4 present: [2, 3, 4, 5, 6, 7, 8]
agent-2 waited 2 mins
agent-2's cost: 0, net reward: 0
agent-2's cumulative reward over 8 proposals: -4
successes to date: 1.000
proposals to date: 8.000
success ratio: 0.125
reward ratio: -0.250

```

Fig. 2: Extract of a Simulation Log

3.2 Results

In analysing the results of Experiment 1, we focus on two dimensions for measuring the outcome: **interval overlap** and **proposal success ratio**. For convenience, we repeat a slightly modified version of Definition 1:¹⁰

Definition 2 *The **overlap between intervals** ι_P, ι_R is the quotient*

$$\frac{|\iota_P \cap \iota_R|}{|\iota_P \cup \iota_R|}$$

Definition 3 *The **success ratio** for an agent is the quotient*

$$\frac{\# \text{ of successful meetings}}{\# \text{ of proposals}}$$

In Fig. 3 we plot the average degree of interval overlap for a population of five agents over 250 runs.¹¹ We illustrate four cases, one where there is no learning, and three where the learning factor λ is set at increasingly high values. Fig. 3(b) shows that even a rather small value for λ is significantly better than no learning, and that the overlap between intervals ends up oscillating between 0.8 and 1.0. Fig. 3(c) shows a situation where complete alignment is achieved. By contrast, the setting of $\lambda = 0.5$ produces an oscillation similar to case Fig. 3(b), with the main difference being that this ‘dynamic stability’ is achieved more rapidly.

In Fig. 4, the outcomes for each agent are plotted separately, using the same four values for λ as in Fig. 3. In Fig. 4(c), it is striking that **agent-1** has much

¹⁰ This is known as the Jaccard index of similarity. We have also experimented with a related measure, Dice’s coefficient, which yields comparable results.

¹¹ One average, each agent engages in 250/5 meeting proposals.

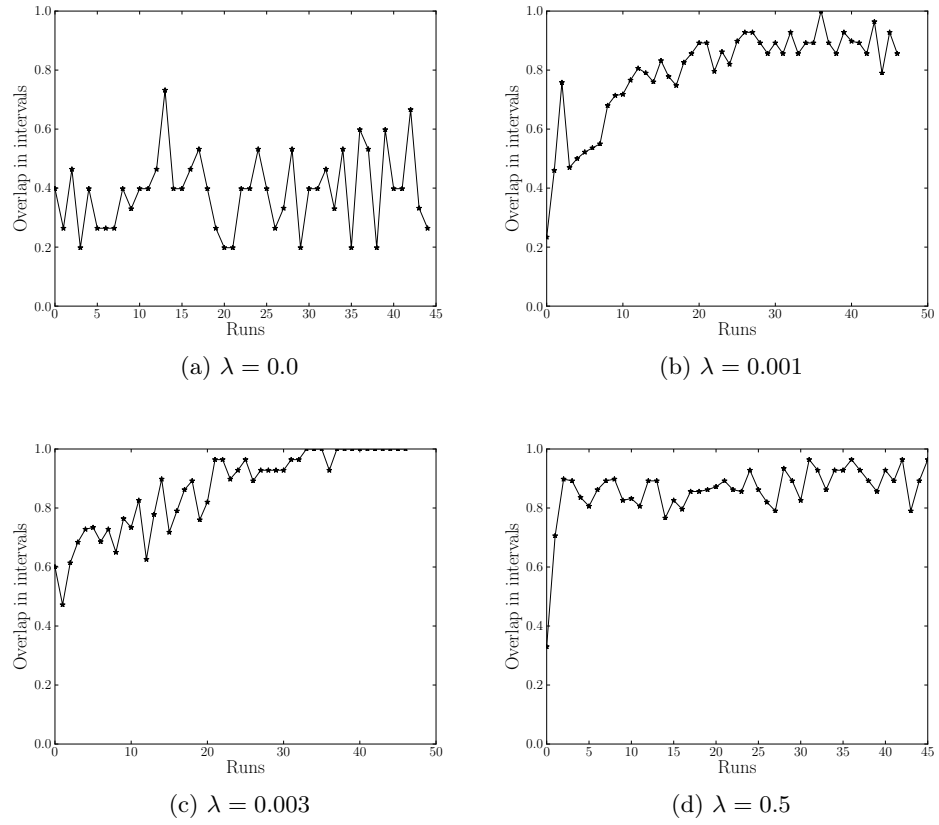


Fig. 3: Average Overlap in Preferred Intervals

lower success than the other agents. This is due to the starting conditions in this particular run, where four of the agents started off with closely overlapping intervals and only **agent-1** happened to diverge sharply from this shared interval.

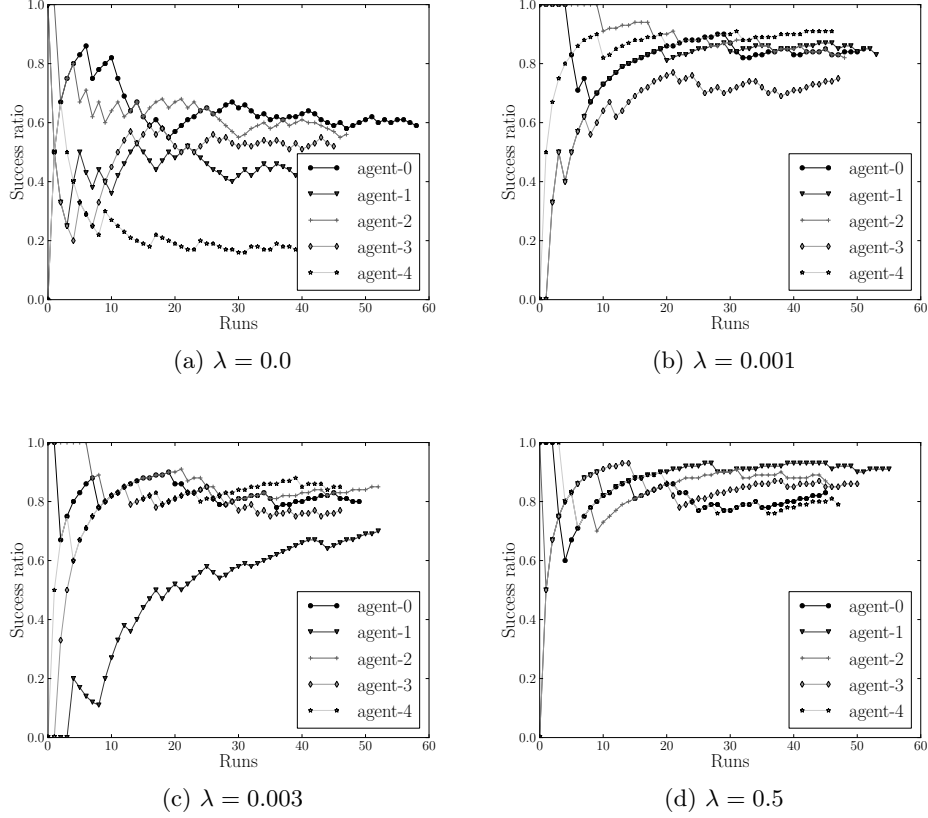


Fig. 4: Ratio of Successful Meetings to All Proposals

Table 1 illustrates the intervals that are associated with each agent at the end of one complete simulation, after alignment has taken place. It can be observed that some of the intervals are left-shifted beyond the earliest point in \mathcal{I} . We will return to this issue later.

3.3 Discussion

As shown in Figure 4(a), when the discrepancy in preferred intervals is allowed to persist throughout the simulation, success in meeting tends to diminish over

agent-0:	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
agent-1:	[-1, 0, 1, 2, 3, 4, 5, 6, 7, 8]
agent-2:	[0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
agent-3:	[-1, 0, 1, 2, 3, 4, 5, 6, 7, 8]
agent-4:	[0, 1, 2, 3, 4, 5, 6, 7, 8, 9]

Table 1: Aligned Intervals after 350 runs, $\lambda = 0.5$

successive runs for all the agents. By contrast, Figure 4(b) shows gradual improvement to a mean success rate of around 0.8 when learning takes place. In addition, a positive value for λ enables the agent population to reach a relatively stable alignment of intervals. Despite this, complete alignment is not typically reached.

Fig. 5 gives an alternative visualization of how the preferred intervals of the five agents become increasingly aligned during the course of successive interactions (under the condition where $\lambda = 0.5$).¹² The greyscale intensity corresponds to the number of agents who share a given time period on a given run: the darker the shade, the greater the degree of overlap across the pool of agents. For example, it can be seen that only one agent has an interval which includes [14, 19] during the first 5 runs, whereas four agents share the subinterval [6, 9]. One way

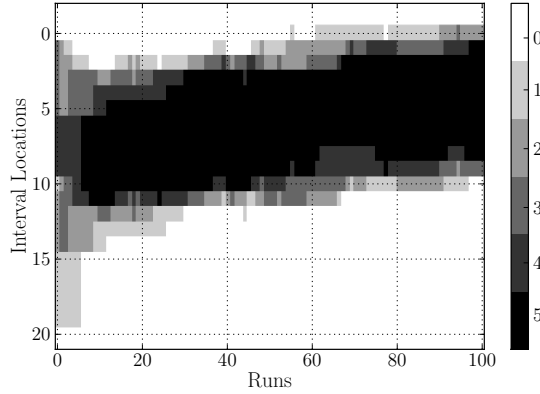


Fig. 5: Distribution of preferred intervals across agents over time

of interpreting this result is to say that the chosen temporal unit is still vague, at the population level, but less so before alignment occurred: the concepts ap-

¹² In order to make this figure more legible, we have truncated the results to only show the first 100 runs.

proximately overlap, in the terminology of Section 1, within the limit set by the error margin ϵ .

If we regard the shared concept as the union of the concepts of the constituent individuals, then we have a *range* of possible boundaries to the concept. This is reminiscent of the ‘egg-yolk’ theory [13, 9] which represents a vague spatial region in terms of its maximal and minimal possible extensions. The maximal extension is called the **egg** and consists of two subregions, the **white** together with the **yolk** (or minimal extension); cf. Figure 6.¹³ If X_t is a snapshot at

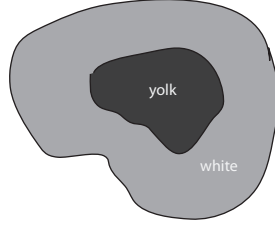


Fig. 6: Egg-Yolk Model of Vague Regions

time t of the vague spatial interval depicted in Fig. 5, then using the Gotts and Cohn [9] predicates `eggof` and `yolkof`, we might try to identify the maximal and minimal regions of X_t in terms of the preferred intervals $\iota_{i,t}$ of agents A_i at time t :

$$\begin{aligned} \text{eggof}(X_t) &= \bigcup_{A_i \in Ag} \iota_{i,t} \\ \text{yolkof}(X_t) &= \bigcap_{A_i \in Ag} \iota_{i,t} \end{aligned}$$

In effect, then, the vagueness is an emergent property of the interaction between the agents in the population, rather than inhering to the conceptual structure of any individual agent. At one level, this perspective has certain attractions, since the indeterminacy of a concept like ‘morning’ does appear to be related to the wide variability in the way that it is applied by individual speakers. Nevertheless, in order to do justice to the intuitions behind the egg-yolk model, we would need to enrich the representation of intervals within agents in order to accommodate something like the egg white region. This would then offer the possibility of agents conditioning their willingness to adapt according to the partition into yolk and white. For example, it might be plausible for a proposer P to only modify its preferred interval if $\mu(\iota_R)$ fell at least within the ‘white’ part of the interval.

One major disadvantage of the framework used in Experiment 1 is that we have, so to speak, ‘hard wired’ the goal of alignment into our agents. It could

¹³ See [3] for a discussion of how the ‘egg-yolk’ model relates to supervaluation [6, 11] approaches to vagueness.

be argued that this has some plausibility; for example, there is considerable empirical evidence that human speakers do align to each other at numerous levels of cognitive representation in dialogue, ranging from phonetics up to the levels of semantic representation and the internal ‘situation model’ [17]. Nevertheless, the process captured in our simulation corresponds more closely to alignment *across* successive dialogues, rather than within a dialogue, which weakens the analogy. Is it possible instead to devise a more principled approach which allows agents to discover the advantages of alignment by themselves?

4 Experiment 2

4.1 Reinforcement Learning

In the second family of experiments, we adopt a simple form of reinforcement learning [23] to replace the alignment strategy of Experiment 1.

Before discussing the details, we need to briefly return to the way in which the proposer P selects a departure time. In Experiment 1, the departure time was set to be the end of the agent’s preferred interval. We now modify this as follows:

$$dep_P = \begin{cases} arr_P + 1 & \text{if } arr_R < arr_P, \\ \text{end of } \iota_P & \text{otherwise} \end{cases} \quad (3)$$

In other words, P departs at time $t+1$ if she knows that R has already arrived (and has either departed already or is still present at time t). Otherwise, P waits until the last point of ι_P . For simplicity, we do not consider the length of the meeting to be a factor in determining costs or utility.

In principle, P incurs a waiting cost which is proportional to the length of the interval $[arr_P, dep_P]$. However, for simplicity, we treat it as a fixed value, regardless of the length of the wait.¹⁴

Let us return to the learning scenario. To ease exposition, suppose that we have a pool of two agents, with a fixed assignment of roles. Each run t of the simulation contains a representation of a state s_t , on the basis of which P selects an action a_t . On the next run, P receives a numerical reward r_{t+1} and finds itself in state s_{t+1} ; the reward is used to build a model of the long-term utility of performing action a in state s_t (taking into account sequences of state changes induced by the action, assuming utility-optimal behaviour thereafter). P maintains a mapping from states to probabilities of selecting each possible action. This mapping is called a *policy*, and is updated in the light of rewards received in states up to and including the current one.

We represent a state with the variable *alignment*. This takes as value one of five possible labels, each of which serves as a bin for a range of integers, corresponding to the difference σ between the median $md(\iota_P)$ of P’s preferred interval and estimated mean $\mu(\iota_R)$ of R’s arrival times. The correspondence

¹⁴ We assume that the cost is zero if $dep_P = arr_P + 1$.

between labels and the value of σ are shown in Table 2. For example, *alignment* would be assigned the value *other_v_early* just in case $md(\iota_P) - \mu(\iota_R) > 6$.

bin labels	range of σ
<i>other_v_early</i>	$\sigma > 6$
<i>other_early</i>	$6 \geq \sigma > 1$
<i>aligned</i>	$1 \geq \sigma > -2$
<i>other_late</i>	$-2 \geq \sigma > -7$
<i>other_v_late</i>	$\sigma \leq -7$

Table 2: Values of the *alignment* variable

The set \mathcal{A} of possible actions for P are analogous to the set of possible alignments:

$$\mathcal{A} = \{\textit{shift_far_earlier}, \textit{shift_earlier}, \textit{no_op}, \textit{shift_later}, \textit{shift_far_later}\}$$

Each action is a mapping from intervals to intervals. The actions *shift_far_earlier* and *shift_far_later* move their input five units earlier or later, respectively, while *shift_earlier* and *shift_later* only move their inputs one unit earlier or later. *no_op* just returns its input unchanged. The actions are defined so that intervals cannot be shifted beyond a stipulated lower and upper boundary (taken to be 1 and 21 in the current model). This constraint is realistic to the extent that, for example, the start point of *morning* would not normally occur before 12.00 am. However, the way that we have implemented these constraints could definitely be improved (for example by defining a probability distribution over possible start times).

Note that despite the potential fit between actions and alignments, any association between the two has to be learned by the agents, rather than being stipulated in the model.

The reward received by an agent depends on the values of two variables *met* and *wait*. The first of these is boolean-valued, while *wait* takes a non-negative integer as value. Rewards are allocated according to the matrix in Table 3. In

	<i>wait</i> = 0	<i>wait</i> > 0
<i>met</i> = True	2	1
<i>met</i> = False	-2	-3

Table 3: Reward Matrix

order to choose an action, the agent estimates the relative values of all members of \mathcal{A} . The estimated value of action a on the t^{th} run in state s is written $Q_t(a, s)$, and we define this to be the average of the rewards received in s by the time the

action was selected. That is, if a has been selected k times in s by the time of run t , giving rise to rewards r_1, r_2, \dots, r_k , then its value is estimated to be the following:¹⁵

$$Q_t(s, a) = \frac{1}{k} \sum_{i=1}^k r_i \quad (4)$$

When $k = 0$, we take $Q_t(s, a) = 0$. $Q_t(s, a)$ is re-computed on each run of the simulation.

The simplest strategy for action selection is the greedy method: choose an action which has the highest estimated value. However, it turns out to be advantageous to behave greedily most of the time while occasionally — with small probability ϵ — selecting at random some other action. We take $\epsilon = 0.1$ initially, and let it decrease over successive runs, so that the action space is sampled more broadly at the beginning of the simulation. This is termed an **ϵ -decreasing** strategy.

The purpose of this approach is to provide a decision-theoretic grounding for the usefulness of alignment. Instead of assuming a hardwired propensity to adjust towards the other agents' concepts, rewards received from the environment alone should be sufficient to cause the agent to behave in such a way, i.e., it would be rational for her to do so, purely on the basis of self-interest.

4.2 Results

In Experiment 1, we only required agents to update their preferred interval with respect to the observed behaviour of their most recent partner. For example, we might have agent A_1 moving ι_i earlier after interacting with A_2 and moving it later on a successive turn after interacting with A_3 . However, in Experiment 2 we also consider the case where agents align not to the pattern of their individual partners, but rather to the mean behaviour of *all* their partners. We shall refer to these two conditions as `ALIGN_TO_GROUP = False` vs. `ALIGN_TO_GROUP = True`, respectively.

Figures 7(a), (b) show the **average reward ratios** achieved over 500 runs. For individual agents, the reward ratio is defined as follows, where K is set to be the maximum possible reward in a state, i.e. $K = 2$:

Definition 4 *The **reward ratio** for an agent is the quotient*

$$\frac{\text{sum of rewards received}}{\# \text{ of proposals} \times K}$$

The average reward ratio is obtained as the mean of the reward ratio taken over the whole population. It remains low throughout the simulation under condition `ALIGN_TO_GROUP = False` (Figure 7(a)), but reaches a point above 0.6

¹⁵ Although reinforcement learning typically involves learning the utility of *sequences* of actions, the more restricted version we have adopted here is sufficient to support our claim that alignment can be learned rather than being hard-wired.

when $\text{ALIGN_TO_GROUP} = \text{True}$ (Figure 7(b)). Analysis of the behaviour of individual agents, illustrated in Figures 8(a), (b), shows that as in the case of Experiment 1, it is possible for one or more agents to persistently diverge from the rest of the group.

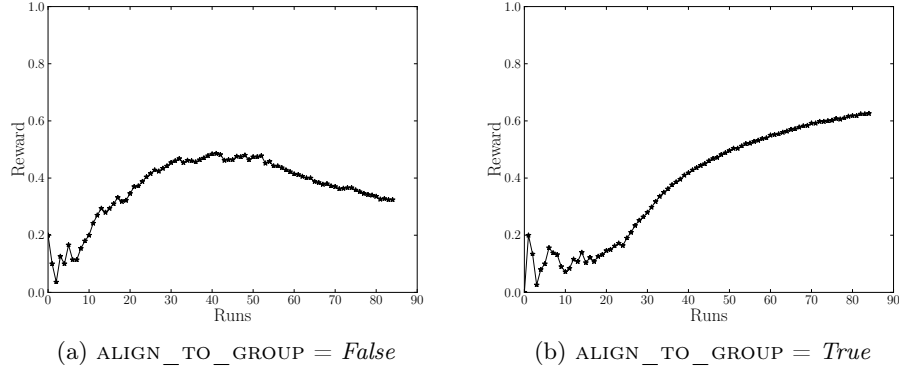


Fig. 7: Average Rewards over Time

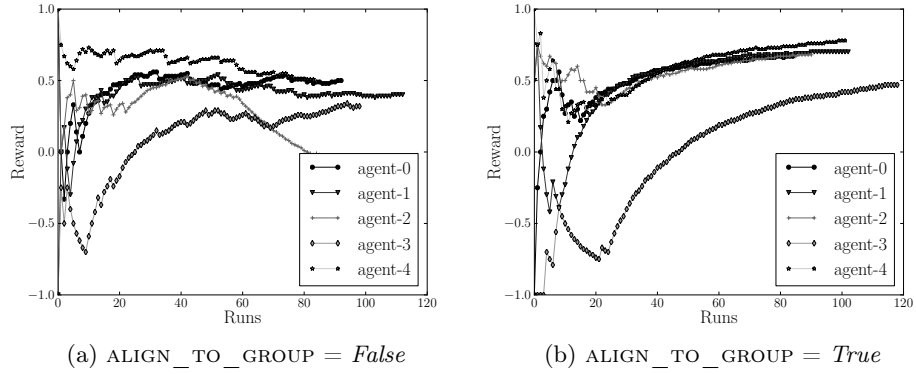


Fig. 8: Rewards to Individual Agents over Time

Figures 9(a), (b) suggest that a reasonably stable alignment of intervals only emerges under condition $\text{ALIGN_TO_GROUP} = \text{True}$.

Finally, Figure 10 shows the extent to which proposals by agents lead to successful meetings.

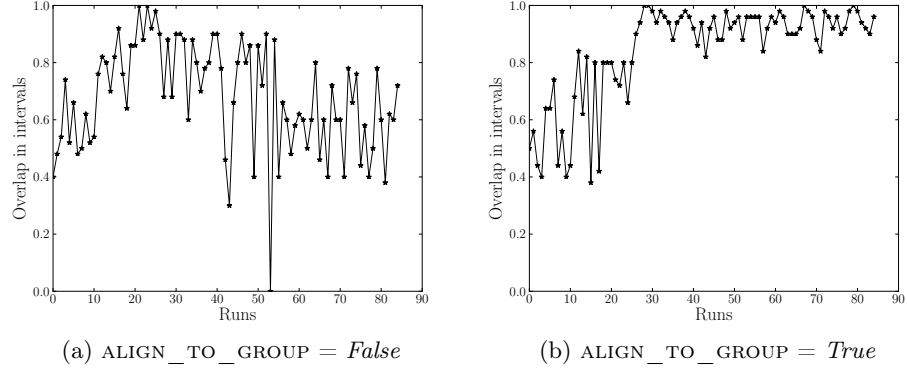


Fig. 9: Average Overlap in Preferred Intervals

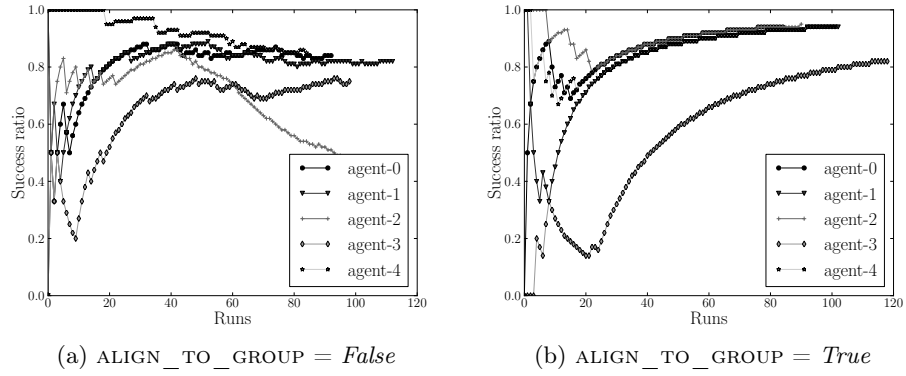


Fig. 10: Ratio of Successful Meetings to All Proposals

In general, the framework using reinforcement learning yields alignment results that are comparable with those achieved with the ‘hard wired alignment’ approach, but only when `ALIGN_TO_GROUP = True`.

These results are comparable to the work on word meaning and multi-agent language games carried out by Steels and colleagues [2, 22, 21]. However, unlike Steels, we are not concerned with how new terms emerge as bearers of meaning, but rather with how pre-existing ‘unstable’ meanings come to stabilize as a result of interaction. In addition, feedback about the interpretation of terms is not acquired through explicit correction and deictic coordination, as for example in [2], but has to be inferred from whether proposals to meet are successfully consummated or not.

5 Conclusions and Future Work

This paper opened with the question *How is it that people manage to communicate even when they implicitly differ on the meaning of the terms they use?* We typically assume that the person we are talking understands our words in just the way in which we ourselves understand them; this is a crucial component of our shared ‘common ground’ [20] in the dialogue. Yet for many items of our core vocabulary, this assumption is probably too strong. Given differences in perceptual apparatus and in personal experience, meanings as mental entities surely differ somewhat from speaker to speaker. Despite these differences, communication usually succeeds, as far as we can tell. We have argued that a notion of **approximate semantic alignment** may be sufficient for communication in a task-oriented scenario. In order to support our claim, we have modelled the utility of a temporal expression for achieving coordinated action, specifically for pairs of agents to arrange meetings between themselves. We have shown that, given certain assumptions, the utility of the expression increases in line with interpretive alignment. That is, when the proposer’s extension for the term overlaps more greatly with that of the responder, then the term is more effective in circumscribing the range of possible meeting times. This in turn increases the likelihood that two agents will successfully meet. If the agents adopt reinforcement learning, then over numerous interactions, they will tend to converge on more tightly aligned sets of interpretations, leading to a stable pattern of successful meeting proposals. However, as we pointed out earlier, our current model only achieves this convergence if agents align to the group average arrival time, rather than successively attempting to align to the average of their immediate partner.

Despite the fact that increased alignment correlates with increased utility, the way we have modelled multi-agent simulation rarely if ever leads to complete alignment. This adds support for the contention that vague terms provide robustness to communication — they work ‘well enough’ in the absence of complete agreement on boundaries. In order to explore this point more fully, let’s return to the details of Experiment 2. Figure 10(b) indicates that one of the five agents (namely **agent-3**) is less successful than all the others — achieving

a score of 0.8 against an average of 0.95 for the other four. Inspection of the simulation log shows that **agent-3** has ended up with a preferred interval that diverges markedly from the rest of the members of the agent pool; see Table 4. Regardless of the reasons why **agent-3** has arrived at a sub-optimal policy, there

agent-0:	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
agent-1:	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
agent-2:	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
agent-3:	[8, 9, 10, 11, 12, 13, 14, 15, 16, 17]
agent-4:	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Table 4: Aligned Intervals after 500 runs, using Reinforcement Learning

is one striking fact: since there is sufficient overlap in preferred intervals, a ‘good enough’ policy can persist. In other words, the residual divergence between intervals across the population does not seriously impede the agents in achieving their goal of meeting.

From a methodological point of view, simulations of the kind presented in this paper do not allow strong conclusions to be drawn, and some kind of analytic model would be desirable. On the other hand, we would argue that simulations do allow us to be explicit about the assumptions we are making and to refine the kind of questions we want to ask. There are a number of issues which we plan to explore in future work, most notably the following:

1. representing temporal intervals using an ‘egg-yolk’ style representation;
2. allowing a responder agent to reject the proposer’s suggestion, and to negotiate an alternative;
3. including in the community certain agents who refuse to adapt to other agents;
4. expanding the range of interactions by providing agents with a lexicon of complementary time expressions (such as *morning*, *midday*, *afternoon* and *evening*);
5. allowing agents to choose temporal expressions at different levels of granularity and to use approximations [12], and
6. applying the approach to the spatial domain.

References

1. Bailin, S., Truszkowski, W.: Ontology negotiation: Ontology negotiation: How agents can really get to know each other. In: Innovative Concepts for Agent-Based Systems. Lecture Notes in Computer Science. Springer (2003) Revised papers from First International Workshop on Radical Agent Concepts WRAC 2002 McLean, VA, USA, January 16-18, 2002.

2. Baillie, J.C., Steels, L.: Shared grounding of event descriptions by autonomous robots. *Robotics and Autonomous Systems* **43**(2-3) (2003) 163–173
3. Bennett, B.: Application of supervaluation semantics to vaguely defined spatial concepts. In: *COSIT 2001: Proceedings of the International Conference on Spatial Information Theory*, London, UK, Springer-Verlag (2001) 108–123
4. Burato, E., Cristiani, M.: Learning as meaning negotiation: A model based on English auction. In Håkansson, A., ed.: *KES-AMSTA 2009*. Number 5559 in *LNAI*. Springer-Verlag (2009) 60–69
5. Euzenat, J., Shvaiko, P.: *Ontology Matching*. Springer, Berlin (2007)
6. Fine, K.: Vagueness, truth, and logic. *Synthese* **30** (1975) 265–300
7. Freyd, J.: Shareability: The social psychology of epistemology. *Cognitive Science* **7**(3) (1983) 191–210
8. Gärdenfors, P.: *Conceptual Spaces: The Geometry of Thought*. A Bradford Book, MIT Press, Cambridge, Mass (2000)
9. Gotts, N., Cohn, A.: A mereological approach to representing spatial vagueness. In: *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR '96)*. (1996) 230–241
10. Jäger, G., van Rooij, R.: Language structure: psychological and social constraints. *Synthese* **159**(1) (2007) 99–130
11. Kamp, H.: Two theories of adjectives. In Keenan, E., ed.: *Formal Semantics of Natural Languages*. Cambridge University Press (1975)
12. Krifka, M.: Approximate interpretation of number words: A case for strategic communication. In Bouma, I.K.G., Zwarts, J., eds.: *Cognitive Foundations of Communication*. Koninklijke Nederlandse Akademie van Wetenschappen (2007) 111–126
13. Lehmann, F., Cohn, A.: The EGG/YOLK reliability hierarchy: Semantic data integration using sorts with prototypes. In: *Proceedings of the third international conference on Information and knowledge management*, ACM (1994) 279
14. McNeill, F., Bundy, A., Walton, C.: Planning from rich ontologies through translation between representations. In: *Proceedings of ICAPS'05 Workshop on The Role of Ontologies in Planning and Scheduling*, Monterey, CA, USA (2005)
15. Ohlbach, H.J.: Calendar logic. In Gabbay, D.M., Reynolds, M.A., Finger, M., eds.: *Temporal Logic: mathematical foundations and computational aspects*. Volume 2. Oxford University Press (2000) 477–573
16. Parikh, R.: Vagueness and utility: the semantics of common nouns. *Linguistics and Philosophy* **17** (1994) 521–535
17. Pickering, M.J., Garrod, S.: Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* **27**(02) (2004) 169–190
18. Reiter, E., Sripada, S., Hunter, J., Yu, J., Davy, I.: Choosing words in computer-generated weather forecasts. *Artificial Intelligence* **167** (2005) 137–169
19. da Silva Sinha, V., Sinha, C., Zinken, J., Sampaio, W.: When time is not space: The social and linguistic construction of time intervals in an Amazonian culture. Accepted for publication in the *Journal of Pragmatics* (to appear)
20. Stalnaker, R.: Assertion'. P. Cole. *Syntax and Semantics 9: Pragmatics* (1979) 315–332
21. Steels, L.: The origins of ontologies and communication conventions in multi-agent systems. *Autonomous Agents and Multi-Agent Systems* (1998)
22. Steels, L., Vogt, P.: Grounding adaptive language games in robotic agents. In: *Proceedings of the Fourth European Conference on Artificial Life*. (1997) 474–482
23. Sutton, R.S., Barto, A.G.: *Reinforcement Learning*. A Bradford Book. MIT Press (1998)

24. Wang, J., Gasser, L.: Mutual online concept learning for multiple agents. In: AA-MAS '02: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems, New York, NY, USA, ACM (2002) 362–369
25. Williamson, T.: Vagueness. Routledge, London, UK (1994)
26. Warglien, M., Gärdenfors, P.: Semantics, conceptual spaces and the meeting of minds. Available from <http://logic.sysu.edu.cn/Article/UploadFiles/200810/20081022091135682.pdf> (2007)