3

Time and space

Contents

3.1 Space 80
3.2 Time 86
3.3 Indexicals, coercion 90
3.4 Measure 93

We owe the recognition of a deep connection between time, space, and gravity to the 20th century, but people have used language to speak about spatial and temporal matters long before the development of Euclidean geometry, let alone general relativity. Throughout this book, we approach problems through language use, in search of a *naive theory* that can be reasonably assumed to underlie human linguistic competence.

Since such a theory predates all scientific advances, there is a great deal of temptation to endow it with some kind of deep mystical significance: if this is what humans are endowed with, this must be the 'true' theory of the domain. Here we not only resist this temptation (in fact we consider the whole idea of linguistics and cognitive science making a contribution e.g. to quantum gravity faintly ridiculous), but we will also steer clear of any attempt to bridge the gap between the naive and the scientific theory. The considerable difference between the two will no doubt have explanatory power when it comes to understanding, and dealing with, the difficulties that students routinely encounter when they try to learn the more sophisticated theories, but we leave this rich, if somewhat anecdotal, field for future study.

In 3.1 we begin with the naive theory of space, a crude version of 3D Euclidean geometry, and in 3.2 we deal with time. The two theories are connected by the use of similar proximities (near/far), similar ego-centered encoding (here/there, before/now/later), and similar use of anaphora (Partee, 1984), but there are no field equations connecting the two, not even in vacuum. The shared underpinnings, in particular the use of indexicals, are discussed in 3.3. Finally, the *naive* theory of numbers and measurement is discussed in 3.4.

A. Kornai, *Vector Semantics*, Cognitive Technologies, https://doi.org/10.1007/978-981-19-5607-2_3

3.1 Space

We conceptualize space from the perspective of the upright human, shown as a cylinder on Fig 3.1. A defining feature of naive space is the *gravity vertical* (Lipshits and McIntyre, 1999). Shown as the dot-dash axis on the figure, gravity is directly sensed by the inner ear, and as such, is a constitutive part of the body schema that we take to be fundamental to the perception of space. Relevant `4lang` definitions include:

```
up       fel         sursum     do_go1ry 763  A
         after(at position), vertical(position er_ gen)
down     le          deorsum    w_do1l1  1498 D
         vertical(gen er_)
vertical fu2ggo3leges verticalis pionowy 869  N
         direction, has top, has middle, has bottom,
         earth pull in direction
fall     zuhan       cado       spadac1  2694 U
         move, after(down)
```

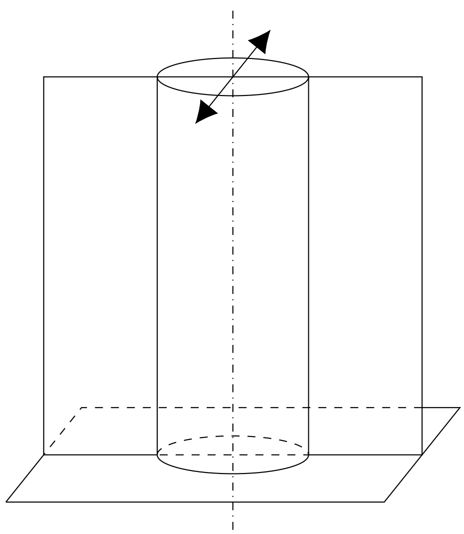

Fig. 3.1: Egocentric coordinates

Orthogonal to the gravity vertical we find a distinguished plane, the `ground`, defined as `surface, solid, at earth`. Actual ground may of course be sloping, but its default orientation is horizontal, as in Fig. 3.1. We use this to define `horizontal`:

```
horizontal vi1zszintes horizontalis horizontalny 3144 A
           direction, flat(ground) has, still(water) has
```

Here we must make the obvious distinction between cognitive *ground*, which we take to be the entire ego-centric model of space and denote by `{place}`, very much including the cylindrical figure at the center, and physical *ground*, which is the flat, horizontal component of this model. What we find under the ground plane is by definition `under` the figure and conversely, the ground plane is defined as the (top plane of the) underside of the schema. One way of saying "if and only if" is to assume that a part of the cognitive schema comes pre-labeled as the underside, and the general relation of `under` is given by the equation

$$P_R(t+1) = P_R(t) + s(|\texttt{=agt}\rangle\langle\texttt{underside}| + |\texttt{=pat}\rangle\langle\texttt{place}|) \qquad (3.1)$$

There is no circularity here. To define spatial notions, we do need an idea of space/place, and we assume this entire model, depicted in Fig. 3.1, to be primitive. (We see this as a prime example of *embodiment*, see S19: Ch.8, but we will not pursue this matter now.) The model has well-defined parts, and these parts are labeled by concepts such as *underside, ground, body,...* For primitives, and only for these, we must take the stance that the concepts are prior (inborn) and language learning consists in attaching names to these inborn concepts.

Next we turn to the cylinder, which we take to be a highly simplified, rotationally symmetric representation of the human body. In effect, the `body` is at the origin of the egocentric coordinate system: a `{place}` always comes with a body at the center, and never mind Cartesian geometry that requires the origin to be a single point with no extension in any direction. Further, the origin already has a definite orientation, it *stands* on the ground. Indeed, while the default of standing is on two feet, we consider it perfectly normal for objects with rough rotational symmetry, such as bottles or vases, to be described as standing on some flat surface.

Even for objects lacking symmetry, such pieces of furniture, it is normal to stand, as long as they have a well-defined top. Objects lacking this feature behave differently, for example it is strange to say that *??the soccer ball stands on the ground*. This is the motivation behind definitions such as

```
stand a1ll sto stac1 74 U
     =agt[vertical], =agt on two(foot)
```

Similar generalizations are available for `foot` and `top`, which we do not at all consider metaphorical in expressions like *foot of the mountain* or *top of the hill*. Some further entries impacted by the egocentric organization of space include

```
base   alap      fundamentum podstawa  146  N
       part_of whole, at bottom, whole has bottom, cause_ whole[fix]
height magassa1g altitudo  wysokos1c1 1583 N
       distance, vertical
root   gyo2ke1r  radix       korzen1   936  N
       under ground, part_of plant, support, at4 base/146
top    teto3     culmen      dach      2377 N
       part, at position, vertical(position er_ part[other])
```

in out The inside of the body is labeled `inside` and the outside is labeled `outside`. This gives our definition of *in* and *out* as

$$P_R(t+1) = P_R(t) + s(|\texttt{=agt}\rangle\langle\texttt{inside}| + |\texttt{=pat}\rangle\langle\texttt{place}|) \qquad (3.2)$$

$$P_R(t+1) = P_R(t) + s(|\texttt{=agt}\rangle\langle\texttt{outside}| + |\texttt{=pat}\rangle\langle\texttt{place}|) \qquad (3.3)$$

When our cognitive ground is a room, we effortlessly identify its ‘skin’ as the walls of the room, its ‘top’ as the ceiling, and its ‘bottom’ as the floor. This actually tells us what’s inside the room and what’s outside of it. We are not bothered by the fact that we can’t identify the arms or the heart of the room, a partial mapping is sufficient for Eqs. (3.2-3) to work as intended.

The next notable feature of Fig. 3.1 is the plane bisecting the cylinder, which we take to be the frontal body plane, given by the maximum extension of the arms. Equally easy to define would be the sagittal plane as the locus of the mirror symmetry the human body enjoys. But we will have much use for arms shortly, whereas symmetries and higher notions of geometry would be hard to justify for the kind of minimalist schema we are developing here. If we permit symmetries, we may as well permit Bessel functions here.

The frontal plane defines the `front` and `back` halfspaces indicated by the two-headed arrow on top. The figure itself provides no clues in this regard, yet most readers will automatically assume that the body is depicted facing the reader, so it is the 7 o’clock arrow that points toward the front, and the 1 o’clock arrow that points toward the back of the body schema. This has to do with a phenomenon that we will discuss in more detail in 3.3: not only do we have our own body schema, one that moves with us as a matter of course, but we also assume that others will have theirs.

No matter how crude an image a cylinder provides for a human body, once we are told that it *is* the image of a body, we start making sense of it in a low-level, automatic fashion. Gordon and Hobbs, 2017 begin their discussion of naive theories with the classic Heider-Simmel test, which shows this phenomenon rather clearly. When we apply the body schema to the human body, it is clear that things `near` are those within arm’s reach (something we could schematize by a larger cylinder around the body), and things `far` are those outside our reach. The space between the internal and the external cylinders can be labeled `about`, and it is only within this space that we can manipulate things (no action at a distance). Relational `about` is anchored to the *about* region of the body in a

about

manner similar to Eqs. 3.1-3.3:

$$P_R(t+1) = P_R(t) + s(|\texttt{=agt}\rangle\langle\texttt{about}| + |\texttt{=pat}\rangle\langle\texttt{place}|) \qquad (3.4)$$

On the figure, it is clear to most readers which is the `left` and which is the `right` side of the body schema. Definitions affected by these notions include most body parts:

```
chin     a1ll     mentum     broda      73   N
         part_of face, at/2744 centre, under mouth
face     arc      vultus     twarz      177  N
         organ, surface, front, part_of head, forehead part_of,
         chin part_of, ear part_of, jaw part_of
forehead homlok   frons      czol1o     1077 N
         part_of face, front, eye under, hair at, at temple/982
front    elej     pars_prior prz1od     608  N
         part, first
nose     orr      nasus      nos        1912 N
         organ, part_of face, animal has face, front, at centre,
         smell, air[move] in, nostril part_of
left     bal      laevus     lewy       222  N
         side, has heart
right    jobb     dextra     prawy      1199 N
         side, lack heart
arm      kar      bracchium  ramie1     1231 N
         organ, long, human has body, body has, limb,
         hand at, wrist at, shoulder at
leg      la1b     pes        noga       1467 N
         limb, animal has, move ins_, support, low
limb     ve1gtag membrum     kon1czyna  3345 N
         part_of body, leg is_a, arm is_a
wrist    csuklo1 articulus   nadgarstek 438  N
         organ, joint, at hand, at end, arm has end
heart    szi1v    cor        serce      2210 N
         organ, cause_[blood[move]], love in/2758, centre
```

That the *heart* is not just the organ of blood circulation, but also the organ of emotions (6.3), love in particular, should come as no surprise: it would be virtually impossible to make sense of much human discourse about love without this assumption. Also, the heart (in Occidental metaphysics, in opposition to the brain, in Oriental, encompassing the brain) is somehow the most central, essential, ruling portion of the body, so that *the heart of the matter* is no more metaphorical than *the top of the hill*.

```
centre ko2ze1ppont centrum    centrum 1412 N
       middle
middle ko2ze1p     media_pars s1rodek 1410 N
       part, place/1026, near centre
side   oldal       latus      strona  1903 N
       part, <two>, centre far, oppose, object has
```

We defer full discussion of some dominantly temporal prepositions such as `follow`, `next`, `(un)til/to` and `through` to 3.2, but we note the strong association to their spatial senses here. By default, people are facing the future and have the past at their back. According to Núñez and Sweetser, 2006, Aymara is an exception, and we see a similar degree of accidentality in linking compass directions to the `place` schema: in Sanskrit `front` is East *pūrva-*, `right` is South *dakṣina*, etc. whereas in Finnish *etelä* 'South' is from *ete-* 'front' and *pohjoinen* 'North' is from *pohja-* 'bottom' (Paul Kiparsky, pc). This situation can be compared to the rule of the road: clearly it makes sense to drive on one side of the road, but it is a matter of convention whether a culture chooses this to be the left or the right side. Importantly, the conceptual schema for compass points may override the 'objective' arrangement of the cardinal directions as in Manhattan, where people will go *North* even when actually they go North-East (see Haviland (2000) for a summary of 'direction keeping' systems and Haugen (1957) for an even more elaborate example).

Of particular interest is the discrete view of space and time imposed by `next`. There are two things involved: a discrete sequence of matters, be they physical objects such as rooms or people standing in a line, abstract entities like numbers or events; and a notion of adjacency among these. When we say *x (is) next (to) y* this means both that x and y are near one another, and that there is no z between them. Typically, this means that x and y are touching, but this is not necessary: we can speak e.g. of *adjacent houses* irrespective of whether garden plots intervene. For an example where contiguity/touching is criterial

`on` consider *on*, which really means 'attached, touching' as in *horseshoes on hooves*. The most frequent (default) case is when the attachment is provided by gravity as in *the book on the table*. This is summarized in the definition `at, =agt touch =pat, <high(=agt er_ =pat)>`.

The notion of `{place}`, as developed in the foregoing, provides our second example of a conceptual schema of the kind words are constantly mapped on (recall `{exchange}` in Fig. 1.2). We could have called this schema *figure-ground complex* or *spatial model* just as easily, but English *place* is quite nebulous (dictionaries from Webster's 3rd to LDOCE list dozens and dozens of meanings) and our technical meaning covers most of these.

The geometry of this voronoid, given to us as as a collection of a few word vectors in L (a space of several hundred dimensions), has nothing to do with the approximate 3D geometry we model in 3.1. The means of guaranteeing that the body axis is aligned with the gravity vertical lie largely outside the domain of linguistic data, in the realm of embodied cognition (Gallese and Lakoff, 2005). When we use *inside* or *body* to label a

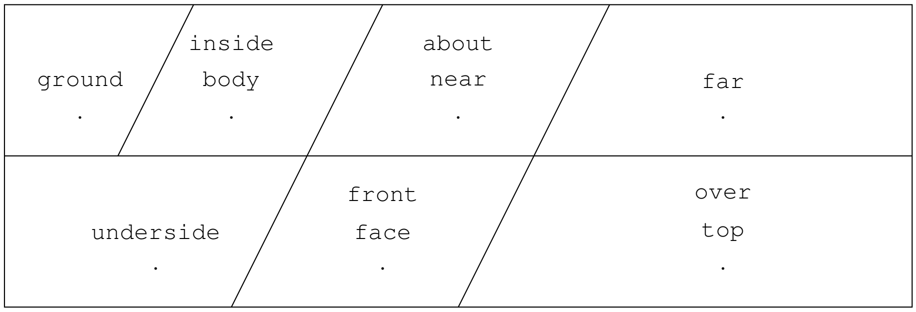

Fig. 3.2: `{place}`

polytope, this means that there is no inside without being inside something, the `body_` segment of the voronoid is automatically invoked.

The relations that obtain between the polytopes are inherent in the schema. To get this effect, we need to go beyond the vectors (polytope regions) depicted in 3.2, and consider the matrices that model two-place relations. Whatever `=agt under =pat` means is derived from the fact that the matrix corresponding to `under` maps the `underside` polytope to the `ground` polytope. In other words, `{place}` is a conjunction of the vectors that make it up, and some canonical equations such as 3.1-3.3 that obtain among these vectors. In fact, we have expressions specifically devoted to signaling major mismatches between these canonical equations (the inherent content of the schema) and a particular situation, e.g. when something is turned *upside down* or *inside out*.

The body schema, sketchy as it is, already provides us with a mechanism to discuss the systematic differentiation that some languages make between intransitives and transitives, (Using the `4lang` notation, we will often speak of U/V alternation.) English often leaves the distinction unmarked: especially in the core vocabulary we find a multitude of examples like *Mary changed* 'she became a different person' versus *Mary changed John* 'he became a different person'. Hungarian offers hundreds of roots that exemplify the intransitive/transitive alternation, and these can generally be translated with 'be(come) X' versus 'make X'. For example, *bús* means 'sad', *búsul* means 'be sad' and *búsít* means 'make someone sad, sadden someone'. Similarly, *but(a)* means 'stupid', *butul* means 'become stupid' and *butít* means 'make someone stupid'. Based on these two examples it may be tempting to think of the root as an adjective, but this is somewhat misleading, as the typical translation (at least to English, Latin, and Polish) is verbal.

In terms of the body schema, whenever the locus of the root X is inside the body, we treat the expression as intransitive, and whenever it is outside, but nearby (within arm's reach) we treat it as transitive. A clear example is *ford* 'turn', where *fordul* means `=agt turn` and *fordít* means `=agt turn =pat`, but the main class of what we called 'mixed `U/V` verbs' in 2.5 also belong here. In the transitive use *John dropped the keys* the locus of the dropping is the object, and in intransitive usage *John dropped* it is the subject, the body itself. We note here that since Fillmore, 1968, having a cognate object is often considered diagnostic of a verbal primitive, as it confers near-root status on the verb by virtue of being identical in intransitive and transitive contexts, even though the latter are limited to cognate objects. (See Höche, 2009 for a detailed discussion of cognate objects in English.)

The `{place}` schema is by no means the only conceptual schema we rely on in conceptualizing space around us, but to complete our discussion of the core spatial vocabulary we need to discuss only one other schema we will call `{bound}`. (Other notions grouped together by Buck (1949) under " Spatial relations: place, form, size" include *change/exchange*, see 3.3; *sign*, see 2.5, and *size*, see 3.4.) The `{bound}` scheme has two spatial participants, the distance, area, or volume that is being bound, which we will call `volume_` irrespective of dimension, and a `boundary_` which typically has one less dimension, e.g. a distance (line, one dimension) is bounded by points (zero dimensional). We could to some extent relate this to the `{place}` schema, comparing the `volume_` to the `body_`, but really the 'skin' that bounds the body is derived from the `boundary_` and not the other way around. Equally important, in a *bound* statement we don't particularly identify with the spatial viewpoint of the `volume_` or the `boundary_`. Rather, the observer is floating somewhere, does not matter where.

A central instance of the schema is provided by *distance* 'the amount of space between two places or things'. In `4lang` we define `distance` as `space/2327 has size, space/2327 between`, and either *distance* directly, or *between*, or perhaps both must make reference to the `{bound}` schema. In one dimension it is clear that the second argument of `between` is a collective noun, composed of two points. In two dimensions, more complex collective nouns are often seen: *[Ann Arbor's] Third Ward, bounded by Huron Parkway, Glacier Way, and US 23, [. . .]*. Remarkably, the boundaries may be only implicitly given, as in *French Guiana is between Suriname, Brazil, and the Atlantic Ocean* rather than *??? between the Suriname-FG border, the Brazil-FG border, and the Atlantic seashore*. The choice between calling the schema `{bound}` or `is_located_between` is rather arbitrary – we were influenced by English *bind*, but note that e.g. in Hungarian the notions of being delimited and being bound are lexically unrelated, and the same is true for Latin *contineo/includo* on the one hand and *astringo* on the other.

3.2 Time

The simplest model of time is the one depicted in the right panel of Fig. 1.3: there are only `now`, and not-now, some `other` time. Sometimes it's light, sometimes dark, sometimes raining, sometimes not, sometimes we hunger, sometimes we are full. In principle, we could consider an even simpler model, depicted in the left panel, where there is only one time, which we can call *now* just as well as we could call it *eternity*. But this doesn't quite amount to a model of time, because we can't have landmarks, we are always in *now* or, what is the same, always in eternity. Ecclesiastes, as good an exposition of the eternal mode of thinking as any, actually relies on a two-state model:

to every thing there is a season . . . a time to be born, and a time to die; a time to plant, and a time to pluck up that which is planted.

The standard mathematical model for now/other is the cyclic group of two elements, C_2, and for n states, C_n, Many temporally marked utterances already make perfect sense in C_2, and not just the ones marked for present tense. What does it mean that *it has rained?* According to Ecc. 1.9 this is no different from *it will rain,* since "the thing that has been, it is that which shall be". Languages like Chinese make no tense distinctions as such in the grammar (this is not at all the same as not making temporal distinctions conceptually), and several languages stop at C_2. For some of these, like Arabic and Japanese, the "other" state is past, so that "now" is lumped together with the future. For others, such as Quechua or Kalaallisut, it is the future, so that the present and the past are lumped together. We will deal here mostly with the C_3 case, with the standard division into past, present, and future, but it is well known that more complex systems exist, by subdividing the past in two (remote and recent) or even three (historic, remote, recent) and similarly for the future.

While cyclic groups are eminently suitable for seconds, minutes (C_{60}), hours (C_{24}), days (C_7), weeks, and in general for calendar devices, we will not spend a great deal of effort on exploring this connection, since calendars are culture- and language-dependent, whereas our focal interest is with universal semantics. For C_2, the conceptual relation between modulo 2 counting and grammatical conceptualization is evident in frequentative forms, which enforce some cyclicity in the way we conceptualize time, but this no longer works for C_3 and higher moduli.

Aside from the particular world-view presented in Ecclesiastes, we consider the past gone, the future unwritten, and the idea of "nothing new under the Sun" that connects these two is uncharacteristic of everyday thought, where no amount of going forward in the future can take us back to the past. One area that makes the weakness of the connection with C_2 evident is iterating *other*, as seen e.g. in the treatment of redoing, which we will briefly inspect in a somewhat underused corner of the lexicon. One meaning of the English prefix *ana-* that we see in Latinate words such as *anabaptism* 'rebaptising'; *anabiosis* 'return to life'; *anaclasis* 'reflect, turn back'; *anacrusis* 'pushing back'; *anadiplosis* 'repetition of a prominent word'; *anaphoric* 'repetition of a word'; *anaphylaxis* 'severe reaction to second or later administration of a substance'; *anatexis* 'melting again'; *anatocism* 'the taking of compound interest' is precisely this redoing. (The most frequent use of *ana-* is in a different sense, 'up', but clearly none of our examples involve this sense.)

Doubling back, returning, does not take us to the original concept, it takes us to the concept *again*, with some temporal marker or counter updated. Getting rebaptized is not at all the same as getting baptized. Extreme reaction to encountering some allergen the second time is sufficiently different from getting this reaction on the first encounter to have its own name. The rocks formed by re-melting and subsequent re-hardening can be distinguished from those formed by hardening alone. Interest on interest adds up, rather than taking us back to the capital. The temporal marker must be (for all cases other than

Ecclesiastes) more complex than C_2, but cannot be as complex as the integers, let alone the reals (see 3.4). Perhaps unsurprisingly, we will reach very similar conclusions in relation to negation in 4.4, where we will conclude that double negation is not the same as affirmation.

`before`

Relative to now, past events are `before`, and future events are `after`. Of these two, a reductionist theory such as `4lang` can treat one as primitive, and define the other based on the observation that

$$x \textit{ before } y \Leftrightarrow y \textit{ after } x \tag{3.5}$$

We will take `before` to be a primitive, irreducible notion, even though there is a tenuous connection with spatial *before*, as in *The knight kneeled before the king*, and we will not take full advantage of Eq. 3.5 because we will have both `before` and `after`

`after`

anchored in `now`, but in slightly different ways. *after* is defined by regular succession, `follow, in order/2739`, and will have many uses in the core vocabulary

`burn`
`stop`

to capture result states, as in *burn* `fire, <=agt wood, after(ash)>` or *stop* `after(=agt lack move)`.

`move`
`exchange`

In contrast, `before` appears mostly in definitions that also have an `after` component, such as *move* `before(=agt at place/1026), after(=agt at other(place/1026))`, or *exchange* `before(=pat at person), after (=pat at other(person))`. By lumping 'before' and 'now' together these could be reduced to pure resultatives, but we will not follow this path here, especially as there

`sudden`
`win`

are other lexical items whose definition refers to pure preconditions: *sudden* is given as `lack warn, before(lack (gen know))` or *win* as `best, succeed/2718, before(compete), before(effort), get/1223 <prize>`.

What the foregoing suggests is that to make sense of temporal effects in terms of world vectors we need not just one world model V, but three: V_b, V_n, V_a 'the world *b*efore the event, *n*ow, and *a*fter'. Aside from the time indexing, we assume these are given to us in the same basis. (In C_2 we would use only two, V_n and V_o.) The key to understanding temporality is that these worlds themselves are timeless, and the time spent between them is underspecified. Silently, automatically, V_n becomes V_b, and V_a becomes V_n, and we need predicates only when objects and their properties do not persist. We only need to list the changes to obtain a full picture of the next time instance based on the present one, and conversely, we generally only find changes to be worthy

`former`

of reporting. A good example is *former*, which simply means that the object is in V_b at the relevant coordinates, but not in V_n. When exactly that former state obtained (a few minutes, days, or many years ago) is something that pure tense marking, such as a PAST morpheme, leaves unspecified, especially in systems where it is not contrasted to historical/remote past marking.

In 1.4 we sketched the `exchange_` frame we invoke in analyzing commercial exchange (buying and selling), as distinct from the word `exchange` which has no commercial aspect, cf. *they exchanged knowing glances*. Now we can extend this to a fully temporal analysis. We have four participants: Buyer, Seller, Money, and Goods, and we assume the exchange itself is taking place now. Therefore, in V_b we have `B has M, S has G` and in V_a we have `B has G, S has M`. This is invariant of the choice of *buy* or *sell* perspective, and more important, it is the most exhaustive statement we can make about the temporal ordering without introducing spurious ambiguities. In reality, the exchange of money may precede, follow, or be synchronous with the exchange of the goods: we don't know, and we conceptualize the whole an act of buying/selling independent of this detail. The frame actually carries no hidden presupposition that normally the goods are handed over first, or the other way around. It's not that the various orders cannot be expressed, they can, but it requires special effort to disambiguate between them.

There are significant differences between the naive view of time presented here and the scientific model. In fact, the differences between the classical Newtonian and the modern relativistic view, significant as they are, pale into insignificance when comparing either of these to the naive view. Crucially, the naive view is based on discrete time instances, whereas the scientific view relies on continuous variables and differentiability. In 6.1 we will return to the question of how much dynamics can be stated commonsensically – here we offer only a few pertinent observations.

First, there is no guarantee that the left difference Δ_{bn} and the right difference Δ_{na} are in any way similar, in fact there is good evidence to the contrary. Consider *pause*: `pause`
under any theory it will mean something like `lack action, before(action), after(action)`. The usual epistemic limitations, which we will turn to in 6.3, apply to any use of the word: when we say *Hearing the extraordinary noise, John paused typing* there is no guarantee that he will resume typing, the noise may just be the building collapsing on him, but our normal expectation is that he will. We defer a more detailed (nonstandard) analysis of pausing to 6.2, but note here that the speed with which the activity is paused, on the order of human reaction time, say 200 milliseconds, can be very different from the timescale on which the activity is resumed, say after a few seconds of contemplation.

Second, there are key cases where we can't even estimate Δ_{bn} and Δ_{na}, only Δ_{ba} is available: consider *move* defined above as `before(=agt at place/1026), after(=agt at other(place/1026))`. Clearly, the naive theory is too weak to support Aristotelian dynamics (where speed, as opposed to acceleration, is proportional to force) let alone Newtonian, for this would require second derivatives where we don't even have first ones. What little dynamics there is follows neither Aristotle's law that things return to their natural 'rest state' nor Newton's law of inertia that things will keep moving as long as counteracting forces are absent. The best we will be able to provide in 6.1 is Buridan's theory of *impetus*, including the scientifically false, but commonsensical idea that planetary orbits are to be explained by circular impetus. This is consonant with every child's expectation that things on a circular path will continue their circular motion.

From a cognitive perspective, this lack of dynamics is as it should be, especially as *move* is applicable in a great number of cases where motion does not involve physical motion at all: consider *the lecture moved from theory to practical issues* or *they were* *moved to tears*. The same phenomenon of using motion verbs where there is no physical, or even emotional, motion, is seen in verbs of fictive motion: *the pipe runs underground, the fence zigzags from here to the house, the mountains surround the village, …* (Talmy, 1983). There are a number of theories addressing these: at one extreme we find Jackendoff, 1983, who denies that motion is taking place, at the other we find Langacker, 1987, who is basing his theory on the motion of the scanning focus of the observer. (While our sympathies are with this latter view, we cannot possibly adjudicate the matter here, and refer the reader to (Waliński, 2018) for further discussion.)

The paucity of testable predictions in regards to dynamics can be contrasted to the richness of grammatical evidence about perspective. As before (S19:6.4), we consider a Reichenbachian view, distinguishing four different notions of time: (i) *speech time*, when the utterance is spoken, (ii) *perspective time*, the vantage point of temporal deixis, (iii) *reference time*, the time that adverbs refer to, and (iv) *event time*, the time during which the named event unfolds. So far, we spoke only about event time, which we can fairly identify with V_n. Temporal adverbials, such as *quick* defined as `act in short(time)` refer to the size of the temporal interval between V_b and V_a. Speech time and perspective time rarely coincide. Even in blow-by-blow descriptions given in the present tense *so I'm walking down the street, minding my own business, when this guy starts shouting in my face and …* we automatically assume a perspective time prior to speech time.

`quick`

Within the confines of this volume we cannot pursue the issue of *Aktionsart* in any detail, but a few remarks are in order. Obviously, the use of `before` and `after` is closely related to lexical aspect, but on our view semelfactives (Comrie, 1976) like *blink* have `before` and `after` clauses. Analogous to our analysis of `pause`, *blink* would be defined as `before(eye[open]), eye[shut], after(eye[open])`. In contrast, statives like *know* and possessive *have* are defined without reference to an `after` state, and their well-known durativity (once you know something, you keep on knowing it, and once you have something, you own it forever) is due to a general law of default continuation (see 6.4). Telic words exhibit a contrast between their their `before` and `after` (goal) states: for example *release* `before(keep), after(free)` or *drown* `before(breathe), after(dead)`.

We maintain temporal deixis by means very similar to the ones used in maintaining spatial deixis, chiefly by *indexical* expressions, to which we now turn.

3.3 Indexicals, coercion

Perhaps the conceptually simplest way to specify *when* and *where* is by means of absolute coordinates: *the party will start at 2PM on July 29 2020 at (47.55625, 18.80125)*. This powerful combination of simplicity and precision is achieved by reliance on highly complex notions, such as the real line, or spherical coordinates, that are relatively recent developments. Natural language has supported locating matters in space and time for many millennia without absolute coordinates. As in most domains, the central method of transferring information is by relative coordinates, tying the new information to something assumed mutually known, as in *The battle took place in the 32nd year of King Darius' reign*.

In between the relative and the absolute mode there lie centuries of standardization efforts gradually moving us from highly subjective measures like *a few hundred steps* or *two day's journey* to the contemporary metric system of units, made ever more precise by metrology (Mohr, Newell, and Taylor, 2016). Most of the units relevant to natural language semantics are, by contemporary standards, highly imprecise: to keep good track of *years* already requires astronomical observations, *seasons* depend of the vagaries of the weather, *days and nights* are not of equal length, what is seven days' journey for one party may only take six days for another, and so on.

Here we follow in the footsteps on Meinong (see in particular Parsons, 1974, Parsons, 1980) and consider words to be capable of denoting objects about which we only have partial information, partial even to the extent their very existence and identity are uncertain. These denotations are greatly similar to the *pegs* of Landman, 1986, especially as we already have a naturally defined partial order on our hands, containment of polytopes in Euclidean space. Since containment is affected by choice of scalar product, things are a bit more complicated than in the data semantic view proposed by Landman, but on the whole we see no need to introduce new, special entities for indexicals.

There are, broadly speaking, two schools of thought: under the dominant view indexicals are variables that obtain their value in reference to external objects present in the real-world context or elsewhere in the discourse. Under the minority view that we follow here, indexicals are just words, not particularly different from other nouns, common or proper, in the degree to which they are underdefined. We can liken them to *bobbers:* much as the float fisher's bobber keeps the bait at a certain fixed depth, bobbers are partially defined individuals already tied to some properties that can be effortlessly computed from regions of the thought vector that can lie outside the linguistic subspace L. When the water level rises, the bobber rises with it, and so does the bait linked to it by a fixed length of string.

The string has zero length with indexicals like *now* – as absolute time moves on, so does *now*. We don't have a full understanding of circadian clocks (the 2017 Nobel prize in medicine was awarded for discoveries of molecular mechanisms controlling the circadian rhythm in fruit flies) but by definition the state of the suprachiasmatic nucleus, and indeed the state of the entire of hypothalamus, is included in Ψ, and we need no special mechanism for *now* to key off of Ψ. With words like *today* the string is longer, and an absolute value cannot be specified without reference to the current time, but a definition `day, now` is sufficient. For *yesterday* we have `day, after(today)`. (This is not a typo: `after` refers to the result state of what happens after the definiendum. After yesterday, today happens.) In terms of the geometric view (1.4) indexicals are simply polytopes whose distinguished point is obtained by projecting the whole thought vector Ψ in the linguistic subspace L discussed in 2.3.

`today`
`yesterday`

In the spatial domain, the zero length case is *I*, computed effortlessly from the real world speech situation based on `person, speak`. As we discussed in 3.1 *here* is egocentrically attached to the origin of the coordinate system of the speaker `at I`, unless accompanied by a pointing gesture as in *we should plant the tree here*. 2nd person singulars are again automatically resolved to the hearer, but 3rd person requires either deixis or some circumlocution, as does *there, then*. In terms of simplicity, we consider the direct deictic reading of indexicals to be prototypical, but there is often an indirect reading, tied e.g. to perspective time rather than speech time, the coordinate system of the protagonist rather than the ego. Consider *Roxanne hasn't seen such enthusiasm for years* – clearly, *such* refers to the enthusiasm she sees at event time.

`I` `here`

Interrogatives (on our analysis, the morpheme *wh*) are simply requests for a resolution. That they can often be satisfied by a pointing gesture goes to show that the answer is typically obtained by a mechanism outside L proper, engaging those parts of the thought vector that are clamped to visual input. This mechanism of going outside one's own linguistic state vector is also responsible for direct manipulation of the listener's thought vector in rhetorical questions, and in the case of informational questions, by reliance on the knowledge state of the listener.

With a rough understanding of indexicals in place, let us now turn to the general mechanism of *coercion*, what (Fauconnier, 1985) calls 'projection mapping'. It is this, as opposed to the more widely used variable binding mechanism, that we make responsible not just for the interpretation of pronouns, but also for most conceptual analysis. We begin with a simple example we already touched on in 1.4, the *commercial transaction* or `exchange_` schema.

There are four participants: the buyer, the seller, the goods, and the money. Of these, the two agentive forms are compositionally named (see 2.1 where the suffix *-er/3627* is discussed), meaning that *buyer* is agent, and so is seller. As we already noted, the name *money* is somewhat imperfect for the 'thing of value' that is used in the exchange, and *goods* is a very imperfect name. Nevertheless, whatever was the patient of the exchange is forced or *coerced* into this role by a rule of English grammar that the NP following the verb is the patient (see Fig. 1.2). Even more remarkably, whatever appears in the fourth slot is now a 'thing of value', even if it's just a bowl of lentil stew.

If this happened in the interpretation of a single sentence we could claim the effect is due to the preposition *for*, but as the story is told (Gen 25:29), Esau is asking for food, and Jacob asks Esau to *sell* his birthright. Esau was only asking for food, and it is Jacob who invokes the exchange schema, with the slots `seller` filled by Esau; `buyer` by Jacob; and `goods` by the birthright. Subsequently the schema is ratified by Esau swearing to it, and fulfilled by his eating the bowl of lentils. That this food is the 'thing of value' is unquestionable, but how does the vectorial semantics reflect this?

The four vectors $\{\mathbf{v}(\textit{buyer}), \mathbf{v}(\textit{seller}), \mathbf{v}(\textit{goods}),\ \mathbf{v}(\textit{money})\}$ are the defining elements of the exchange schema as a set (we use curly braces to emphasize that their order is immaterial). Together, they define a polytope, the intersection of the positive half-spaces. The other 4 vectors in our example, $\mathbf{v}(\textit{Jacob}), \mathbf{v}(\textit{Esau}), \mathbf{v}(\textit{birthright})$

and $\mathbf{v}(\mathit{food})$ are just points (or small polytopes) in L. What we are looking for is an equaliser Q such that after applying Q to the representation space R that reflects the state of affairs before Jacob making the offer, we obtain R', where not just $\mathbf{v}(\mathit{buyer}) = \mathbf{v}(\mathit{Jacob}), \mathbf{v}(\mathit{seller}) = \mathbf{v}(\mathit{Esau})$ and $\mathbf{v}(\mathit{goods}) = \mathbf{v}(\mathit{birthright})$ but also $\mathbf{v}(\mathit{thing\_of\_value}) = \mathbf{v}(\mathit{bowl\_of\_lentils})$ holds.

These equations are created by several different mechanisms. The first two equations come from resolving pronouns corresponding to speaker and hearer: the sentence *Sell me (this day) your birthright* is addressed to Esau, making him the seller, and spoken by Jacob, making him the buyer. Since the birthright appears in the patient slot of this sentence, we obtain the third equation by the same syntax mechanism we discussed in S19:5.3. The last equation is supported by the mechanism of pragmatic inference discussed in S19:5.7: we know from earlier sentences that the *food* and the *bowl of lentils* are the same, we know Esau is faint, and he himself acknowledges that at this point the food is more important to him than his birthright: *Behold, I am at the point to die: and what profit shall this birthright do to me?* This establishes, from the seller's perspective, that the thing of value to be received for the goods is indeed the food.

By change of scalar product, the vectors corresponding to the discourse entities can be easily moved to the respective positive half-spaces, as discussed in 2.3. But here we want to express not just containment. `Esau is_a seller`, but rather equality, that Esau is (uniquely) tied to the seller slot in this particular instance of the exchange schema, hence the need for equalizers. It adds to the challenge that `exchange_` is not a word, something we could describe by a single vector: we need four vectors to make sense of it, and we know that a great deal of additional knowledge is implicated, such as the reversal of ownership of both money and goods on completion of the schema.

In general, none of the relationals we discussed in 2.4 has a clear and unambiguous word we could use to name it, with the possible exception of `er_`, which has a good, albeit morphological rather than word-level, expression in the English comparative suffix `-er/14`. These relations (the list includes not just spatials but also `cause_, for_, has, ins_, lack, mark_`, and `part_of`) have in common the obvious requirement of using at least two vectors to characterize a single instance, but otherwise they are rather different. More detailed analyses are provided for `has` in 2.2; `for_, ins_` and `part_of` in 2.4; `mark_` in 2.5; `lack` in Chapter 4; and `cause_` in 6.2.

3.4 Measure

Counting and measuring things is central to civilization. Buck (1949) lists "Quantity and number" as one of the semantic fields he uses to organize the IndoEuropean material, containing entries not just for the cardinals *one, two,* ... and the ordinals *first, second,* ... and fractions *half, third,* ... but also for less specific notions of quantity such as *much, many, more, little, few* and for broad measure phrases like *full, empty, whole, enough, every, all,* .... As Buck (13.31) notes, "no class of words, not even those denoting family relationship, has been so persistent as the numerals in retaining the inherited words". Given the semantic coherence of the class, and the difficulty of subtle shifts in meaning, it is not surprising that this phenomenon is not limited to IndoEuropean – similar coherence is seen e.g. in Bantu, now tentatively extending to Niger-Congo (Pozdniakov, 2018).

From the mathematical perspective, the first thing to note about the system is that there is no system. It is only in hindsight, from the vantage point of the modern system of natural numbers $\mathbb{N}$, that we see the elements of counting, the cardinals, as being `last` useful as ordinals as well. But certain notions like *last*, '`part_of sequence, at end`' which make eminent sense among ordinals, have no counterpart among the car- `first` dinals, while others, like *first* '`lack before, second/1569 follow`', do. Key elements, *one* in particular, are used not just for counting and ordering, but also for signifying uniqueness 'unus, unicus' and separateness, standing alone.

The idea of using functions from objects to $\mathbb{R}$ to gain traction of measure phrases such as *three liters of milk* is common in mainstream logical semantics (Landman, 2004; Borschev and Partee, 2014) but, as will be discussed in greater detail in 4.5, we view this approach as highly problematic both in terms of empirical coverage and in terms of bringing in an extra computational stratum. `4lang` has no problems handling vague `many` measures of quantity, like *many* '`quantity, er_ gen`' or *few* '`amount(gen` `few` `er_)`', though these present the modern, more precise, theory with significant difficulties. However, it does have problems with the modern quantificational readings of *all* `all` and *every*, since it defines the former as '`gen, whole`' and the latter as `gen`. As we `every` have noted elsewhere (Kornai, 2010b), actual English usage (in newspaper text) is characterized by generic readings, and the episodic readings are actually restricted to highly technical prose of the kind found in calculus textbooks.

Thanks to the foundational work of the late 19th and early 20th centuries we now have a simple, elegant method for extending $\mathbb{N}$ to the rationals $\mathbb{Q}$. These, or even finite precision decimals, would arguably be sufficient for covering much of everyday experience, especially ordinary measure phrases like *This screen is 70" wide*. Since the Message Understanding Conferences (Grishman and Sundheim, 1996) special attention is paid to the extraction of *numerical expressions* (NUMEX) such as monetary sums and dates. The notion of calendar dates has been extended to cover more complex time expressions (TIMEX), and for most of these, there is a standard Semantic Web representation scheme, ISO TimeML associated to instances, intervals, etc. which grew out of earlier work on providing semantics for time expressions (Pustejovsky et al., 2003; Hobbs and Pan, 2004). Extracting this information from (English) text is difficult (Chang and Manning, 2012) and the parsing and normalization of time/date expressions is still an active research area (Laparra et al., 2018).

These representation schemas, both for direct time and space measurements, and for the more abstract quantities like monetary sums, implicitly rely on the standard theory of the real line $\mathbb{R}$. Tellingly, all work on the subject has an important caveat (Hobbs and Pan, 2004):

> In natural language, a very important class of temporal expressions are inherently vague. Included in this category are such terms as *soon, recently, late*, and *a little while*. These require an underlying theory of vagueness, and in any case are probably not immediately critical for the Semantic Web. (This area will be postponed for a little while).

Here we turn this around, and treat expressions like *soon* 'a short time after <now>' or *late* 'after the time that was expected, agreed, or arranged' as normal, and vague only from the vantage point of the arbitrary precision semantics imposed by using real numbers. From this vantage point, every term we use in ordinary language is vague: for example *water* does not precisely demarcate how many milligrams/liter mineral content it may have. From the vantage point of ordinary language, it is not just the real numbers $\mathbb{R}$ that require special semantics, the problem is already present for natural numbers $\mathbb{N}$: iterative application of the Peano Axioms (or even the axioms of the weaker system known as Robinson's Q) is not feasible given the simple principle of non-counting that we have argued for in Kornai (2010b):

> For any natural language N, if $\alpha p^n \beta \in N$ for $n > 4, \alpha p^{n+1} \beta \in N$ and has the same meaning

Since we simply can't make a distinction between *great-great–great-great-great-great-great-grandfather* and *great-great-great-great-great-grandfather* unless we start counting on our fingers, we conclude that the only feasible approach is to do the work outside `4lang` by means of a separate equation solver. This is the approach taken in modern systems aiming at word problems such as Kushman et al. (2014), which derives the equations from text using standard NLP methods, and solves them by Maxima.

Unlike ordinary language understanding, solving word problems, or even setting up the equations, is a skill that Kahneman (2011) would consider 'slow thinking'. Whereas ordinary semantic capabilities are 'fast thinking', deployed in real time, and acquired in everyday contexts by all cognitively unimpaired people early on, solving word problems is a task that many fail to master even after years of formal schooling.

Once we permit an external solver, there is no need to restrict the system to (finite precision) rationals, and sophisticated methods using reals $\mathbb{R}$ and even complex numbers $\mathbb{C}$ are also within scope. What we need is a system to extract the equations from the running text. This is effectively a template filling task, originally considered over a fixed predetermined range of templates by Kushman et al. (2014), and more recently extended to arbitrary expression trees by Roy and Roth (2016). This is a very active area of research, and we single out Mitra and Baral (2016) and Matsuzaki et al. (2017) as particularly relevant for the linguistic issues of assigning variables to the phrases used in the question and in the body of the word problem.

Altogether, the proto-arithmetic that is discernible in systems of numerical symbols, e.g. Chinese 一, 二, 三 or Roman I, II, III or from reconstructed proto-forms that give 7 as '5 + 2' or 8 as '4 · 2' is haphazard, weak, and both theoretically and practically inadequate. This is evident not just from comparing the axiomatic foundations of arithmetic to that of `4lang` but also from evolutionary considerations, as the modern system of Arabic numerals has displaced all earlier ones such as the Babylonian, Chinese, Roman and Maya numeral systems.

It doesn't follow that every semantic field will require a specific, highly tailored system of Knowledge Representation and Reasoning to get closer to human performance, but certainly 'slow thinking' fields will. Such systems actually have great intrinsic interest: for example Roy and Roth (2017) offer a domain-specific version of type theory (better known in physics as dimensional analysis) to increase performance, a deep domain model on its own right. But our interest here is with precisely the kind of 'fast thinking' that does not require deep domain models. We return to the matter in Chapter 8, where we will discuss a central case, trivia questions, which we can capture without custom-built inferencing.

size
dimension

To elucidate the 'fast thinking' theory of quantity further, we consider the notion of *size*, which `4lang` defines simply as `me1ret magnitudo rozmiar 1605 c N dimension`. In turn, *dimension* is `dimenzio1 dimensio wymiar 3355 c N quantity, size, place/2326 has`. We again see a near-mutual defining relation, but with the added information that `dimension`, and by implication, `size` is a quantity, one that `place/2326 has`. Tracking this further, *place/2326*

place/2326
place/1026

is given as `te1r spatium przestrzen1 2326 c N thing in`, related to the `{bound}` schema we discussed in 3.1, as opposed to *place/1026* `hely locus miejsce 1026 c N point, gen at`, which is related the the `{place}` schema. It may be possible to unify these two schemas e.g. by assuming that the body used in `place/1026` is also a `place/2326 has`, but we see no compelling reason to do so, especially as this would bring in the human size scale as default to both, a step of dubious utility.

big
small
large
little
-est
all

Our treatment of measure is geared toward raw measurements, as in *John is tall* or *It was a huge success*, as opposed to measure phrases like *John is six feet two inches tall* or *The earthquake measured 7.1 on the Richter scale*. Raw measurements are treated as comparisons with averages: *big* is defined as `nagy magnus duz1y 1744 e A er_ gen`, and *small* as `kis parvus mal1y 1356 u A gen er_`. (*large* is defined as `big`, and *little* as `small`.) This yields a three-pont scale: big/large – medium – small/little, which can be extended to five points by adding superlatives, typically by means of the suffix *-est*, defined as `leg-bb -issimus naj-szy 1513 e G er_ all`. Here `all` is not some new quantifier, but simply another noun, `mind omnis wszyscy 1695 u N gen, whole`. We defer a fuller discussion of quantifiers to 4.5, but note here that `4lang` treats them as more related to pronouns than to VBTOs.

In Chapter 5 we will use an even finer, seven-point scale to describe the naive theory of probability, but this should not obscure the plain fact that no n-point scale, for however large n, can capture modern usage, which relies on real-valued (continuous) measure phrases, for which we must rely on equation solvers we see as entirely external to natural language semantics. To deal with *six feet two inches tall* we would need some mechanism that shows this to be equivalent to *188 centimeters tall*. This requires not just the foot/inch and inch/cm conversion, but also the capability to recognize that for practical purposes the unrounded value of 187.96 must be rounded. We can't measure people's height to a millimeter, but if we are talking about a uranium rod in a nuclear power plant, we may well insist on this, if only to guarantee that it will fit some precision-manufactured container.

This is not to say that we cannot write a grammar capable of recognizing the measure phrase. To the contrary, building such a grammar is near trivial when the measurement unit is explicit in the text (but can lead to expensive mistakes when it is not), and standard rule-to-rule compositional semantics, specified in terms of ordinary arithmetic operations, can be used to compute values to arbitrary precision. But doing so is irrelevant to our main goal, which is to characterize human semantic competence, rather than the competence of ALUs.