Representing Time in Scientific Diagrams

William Bechtel (bechtel@ucsd.edu)
Daniel Burnston (dburnston@ucsd.edu)
Benjamin Sheredos (bsheredos@ucsd.edu)
Department of Philosophy and Center for Circadian Biology, University of California, San Diego
La Jolla, CA, 92093-0119 USA

Adele Abrahamsen (aabrahamsen@ucsd.edu)
Center for Research in Language and Center for Circadian Biology, University of California,
San Diego, La Jolla, CA 92093 USA

Abstract

Cognitive scientists have shown increased interest in diagrams in recent years, but most of the focus has been on spatial representation, not conventions for representing time. We explore a variety of ways in which time is represented in diagrams by one research community: scientists investigating circadian rhythms at the behavioral and molecular levels. Diagrams that relate other variables to time or indicate a mechanism's states across time use one or two spatial dimensions or circles to represent time and sometimes include explicit time markers (e.g., the hours on a clockface).

Keywords: Circadian rhythms; diagrams; mechanistic explanation; time

Introduction

A number of cognitive scientists have become interested in the interaction between human reasoning and external visualizations. Projects in such areas as knowledge representation, human-computer interaction, and situated cognition have all focused on how information can be represented in a range of distinct formats and used as reasoning tools. Experimental and theoretical work on diagrams in particular has made great strides in recent years (Cheng, 2002, 2011; Gooding, 2010; Hegarty, 2004, 2011; Nersessian, 2008; Tversky, 2011). Still, significant challenges remain in understanding visualization. Our focus is on how diagrams support reasoning in complex empirical domains (Sheredos, Burnston, Abrahamsen, & Bechtel, 2013). A critical challenge researchers face in developing diagrams is how to represent multiple aspects of a problem space. For instance, while two-dimensional diagrams readily support spatial reasoning tasks, many tasks require reasoning about time, and representing time and integrating both spatial and temporal information pose special challenges.

Our strategy in this paper is to examine published diagrams from a field in empirical science that has dedicated significant attention to ways of representing events in time: chronobiology, the study of circadian and other biological rhythms. What is learned here has broader implications.

The term *diagram* is used in both inclusive and restricted senses. In its inclusive sense, indicated by the etymology of the word, *diagrams* are visuospatial representations. All the figures in a scientific paper, including line graphs, typically count as diagrams. Sometimes the term is used more restrictively for graphical representations of the parts and operations of a mechanism. We refer to these as *mechanism diagrams*, and they are of particular interest as they play crucial roles in developing, evaluating, and presenting mechanistic explanations. Biologists often begin by identifying a system that in relevant conditions generates a phenomenon of interest and then seek a mechanistic account of how it does so. This involves identifying its parts, determining the operations they perform, and showing how, when organized appropriately, the parts and operations generate the phenomenon of interest (Bechtel & Abrahamsen, 2005; Bechtel & Richardson, 1993/2010; Machamer, Darden, & Craver, 2000). This practice is often supported by mechanism diagrams in which icons or glyphs (Tversky, 2011) specify parts of the mechanism and arrows indicate the operations by which parts affect other parts or are transformed into other types of parts. However, these mechanism diagrams do not stand alone. To relate parts and operations represented in the diagram to a phenomenon, researchers need to represent both how the phenomenon is realized in time and how the mechanism operates in time. We will examine both.

Circadian rhythms are approximately 24-hour oscillations generated endogenously within organisms that regulate a host of physiological, behavioral, and cognitive functions. They are found in organisms ranging from bacteria and fungi to plants and animals. Much early research focused on the phenomenon of circadian rhythmicity as observed in animals' fluctuating levels of activity. During the last few decades of the 20th century, circadian researchers began tracing these rhythms to intracellular molecular mechanisms involving feedback relations between proteins and the genes from which they are transcribed and translated.

Challenged to understand how individual cells maintain an approximately 24-hour oscillation and how populations of cells synchronize their activity, circadian rhythm researchers have developed a variety of diagram formats. Most straightforward is to map time to one of the two spatial dimensions (or hours to one dimension and days to the other), but this comes at the cost of pre-empting a resource and hence limiting what else can be included. If, a circle is used instead to represent a 24-hour duration, that opens up several ways to incorporate other kinds of information. We will display and discuss examples of how these formats display timing either of a phenomenon or of an operation with-

in a mechanism. We turn in the last section to mechanism diagrams that use both spatial dimensions to represent the parts and operations of a mechanism, and consider techniques that nonetheless can incorporate changes of state in the depicted mechanism over time.

Representing Time on One or More Dimensions

The most straightforward way to represent time in a diagram is to dedicate one spatial dimension to time. Often a bar or line graph is set up with time on the abscissa and a dependent variable from an experiment on the ordinate. When the phenomenon of interest involves oscillation in time—as it often does in circadian research—such graphs will display this as an oscillatory pattern in space (left to right). In Figure 1, for example, Hardin, Hall, and Rosbash (1990) used a line graph to display the repeated rise and fall of the relative abundance of mRNA due to transcription activity of the clock gene *period* in fruit flies. It can be seen that this oscillation across five days has a period of approximately 24 hours. These particular data came from flies kept in constant darkness, demonstrating a key circadian phenomenon: that the daily oscillation is endogenous.

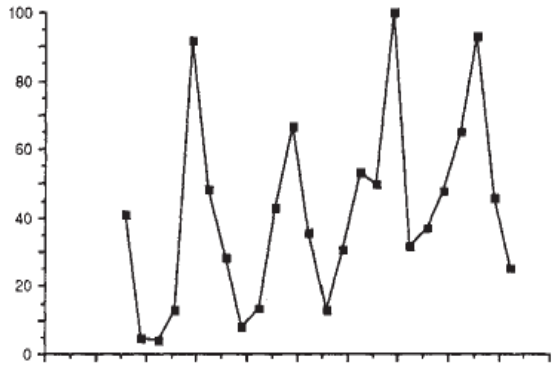

Figure 2. Hardin et al.'s (1990) line graph of changes over a 120-hour period in the relative abundance of *per* mRNA in fruit flies kept in constant darkness.

But are oscillations in darkness exactly the same as those under a normal day-night cycle? Although the line graph offers a direct, visually compelling display of the overall phenomenon of circadian oscillation, using it to address this question would require a close reading of the relevant data points, from which each day's period would be calculated for comparison. Instead, circadian researchers adopted a representational format—the raster plot—that makes comparison across successive days visually accessible. When used to display an organism's activity (rather than molecular concentrations), as in the top panel of Figure 2, such a plot is called an *actogram*. Here both spatial dimensions represent time, but on different scales. Time within each day proceeds horizontally, as in the line graph, but successive days are stacked vertically. Activity at a given time on a given day is indicated by a hash mark and lack of activity by white space. In this example, each horizontal row displays activity from two days (48 hours) rather than a single day, with the second day's activity re-plotted as the left half of the next row. Viewers can choose to focus on the left side to track activity across days, but can also view the entire actogram to better detect any patterns of activity that straddle midnight.

Figure 2 specifically addresses the question of whether oscillations in the running-wheel activity of a mouse are the same in constant darkness as in normal conditions. It shows, visually, that endogenous oscillations get *entrained* by the external *Zeitgeber* ("time-giver") of the day-night cycle, nudging the observed oscillation more precisely to 24 hours. Here is how. A bar at the top indicates which hours this nocturnal animal was exposed to light (white) vs. dark (black)—but only for the first few days (labeled LD). It can be seen that the mouse maintained a precise 24-hour cycle of activity across those days. Thereafter it was placed in constant darkness (DD), and this revealed a free-running period that was slightly less than 24 hours. Activity thus began a bit earlier each day, which shows up in the actogram as a distinctive diagonal pattern. To further explore the impact of external cues, on just one of the dark days the researchers delivered a light pulse (arrow labeled LP) at the time activity would have begun. The sudden rightward shift of the activity pattern demonstrates that a pulse of light is sufficient to *reset* the start of activity (by delaying it several hours), after which the free running pattern resumes.

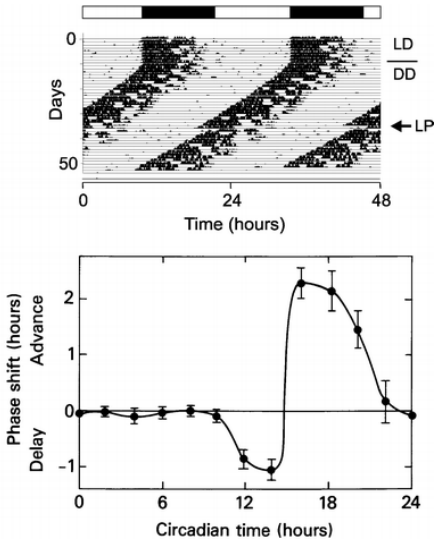

Figure 1. Actogram (top) and phase response curve (bottom) from Lowrey and Takahashi (2004).

Thus, the top panel of Figure 2 makes visually apparent (a) entrainment to the day-night cycle (activity aligned to the light-dark bar, the same in each row); (b) the slightly shorter endogenous period revealed by constant darkness (activity beginning earlier each day); and (c) the resetting of the oscillation's phase—but not of its period—by a pulse of light (rightward shift of activity onset on the day of the light pulse). This last phenomenon can be further explored by manipulating the time at which a light pulse is delivered and then visualizing the effects of pulse timing on phase. A different diagram format is especially suitable for revealing these effects: the *phase response curve*. As shown in the bottom panel of Figure 2, this is a specialized line graph in

which the ordinate indicates by how much (within a 3-hour window) the phase is advanced or delayed, depending on the time of the light pulse—construed here on the abscissa not as clock time, but as *circadian time*, in which a day is based on the free-running period and an hour is 1/24 of that period. This brings an extra complexity in representing time, requiring chronobiologists to make additional inferences to determine the clock time at which a pulse will have the effect plotted. Students can struggle to make these inferences.

An alternative to using hash marks, as in the actogram, is using color to indicate activity. Typically cold colors are used for low activity and warm colors for high activity; accordingly these are often referred to as *heat maps*. Figure 3 presents two heat maps from Ueda (2007), which compare a normalized measure of the expression of 101 genes in the mammalian suprachiasmatic nucleus (SCN) under light-dark (LD) vs. total darkness (DD) conditions. (The gray regions of the DD bar indicate times that normally would have light, but are dark in this condition.) Each horizontal line shows the expression activity of a different gene, and the genes are placed in order of their time of maximum activity (red). Visually, the heat map makes it obvious that (a) the expression of each gene oscillates even without external light cues and (b) there are different populations of genes active during different parts of the day. Thus, circadian researchers have developed a variety of graphical conventions for conveying non-temporal measures when one or both spatial dimensions are preempted for representing time.

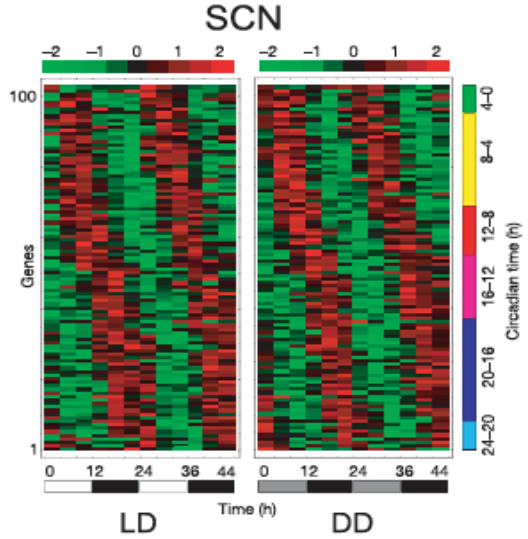

Figure 3. Ueda's (2007) heat maps showing levels of expression of 101 SCN genes in light-dark vs. total dark conditions.

Representing Time on a Circle

Since rhythmic activities regularly return the system to the same state, a circle offers an alternative way to represent time. This, of course, is the representational format that has long been used in mechanical clocks, albeit typically using the circle to represent 12 rather than 24 hours.

Rayleigh plots, such as shown in the bottom panel of Figure 4, illustrate this strategy. Ciarleglio et al. (2009) used a fluorescent marker (GFP) to report expression of the clock gene *Per1* in the SCN, regarded as the central mammalian clock. Data were obtained from mice with normal VIP genes ($VIP^{+/+}$) or lacking one ($VIP^{+/-}$) or both ($VIP^{-/-}$) copies of the gene. The top panel shows the oscillation in *Per1* expression in numerous SCN neurons for each condition. The loss of synchrony in the $VIP^{-/-}$ mutant is apparent, but the Rayleigh plots at the bottom make this even clearer by abstracting away from the detail in the line graphs to focus only on the time at which *Per1* expression reaches 50% of maximum for each neuron (each indicated by a blue arrowhead on a 24-hour clockface—clearly much less clustered in the null mutant). The data also are analyzed statistically to characterize synchronization: the red arrow in each Rayleigh plot points to the mean time expression reaches 50% of maximum (its mean phase), and its length is inversely proportional to the standard deviation. The very short arrow in the right panel indicates both the change in the mean phase and that the phases are much more variable across individual neurons.

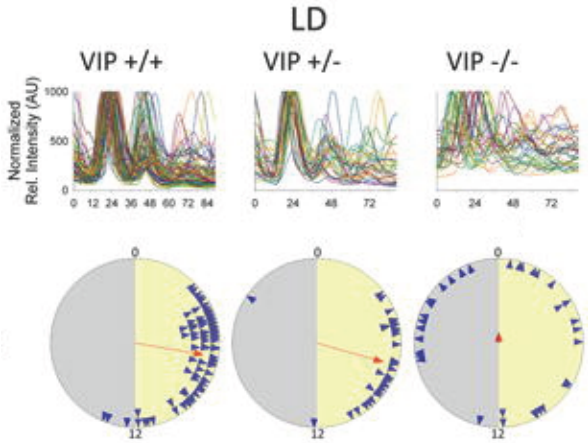

Figure 4. Ciarlegio et al.'s (2009) line graphs showing oscillations of *Per1* expression in neurons of normal and mutant mice (top) and Rayleigh plots highlighting phase (bottom).

Figure 5 shows a different way to use a clock face. Relógio (2011) surrounded theirs with color-coded concentric rings, each tracking concentrations of one protein to

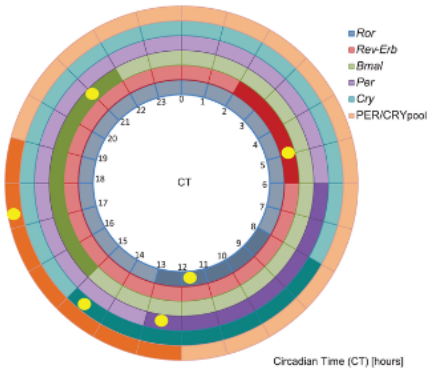

Figure 5. Relógio et al.'s (2011) use of concentric circles around a clock face to represent concentrations of clock proteins across 24 hours.

indicate expression of the corresponding clock gene. The darkened region of each ring indicates when expression of that gene is relatively high in experimental studies, while the small yellow circle indicates time of peak value in a simulation. This visualization makes it evident that the variables reached maximum values in the simulation within but towards the end of the phase of peak expression determined by experimental studies, making clear the model's fit to the data.

Representing Time in a Mechanism Diagram

We turn now to diagrams of mechanisms, in particular, those proposed to explain such relations as are diagrammed in Figures 1-5. Figure 6 is a fairly typical mechanism diagram; it shows the parts, operations and organization of the molecular mechanism thought to be responsible for circadian rhythms in mammals. Parts are represented by various glyphs (colored boxes for genes, correspondingly colored circles and ovals for proteins, white rectangles for promoters, etc.). Arrows show that the operation of a part has an effect on another part. Dark lines indicate spatial compartments (the cytoplasm and the nucleus). Incorporating time is a challenge. The various activities occur at different times of day. The operation associated with each arrow takes time, and each cycle of activity that returns the mechanism to the same state takes approximately 24 hours, but nothing else about timing is shown.

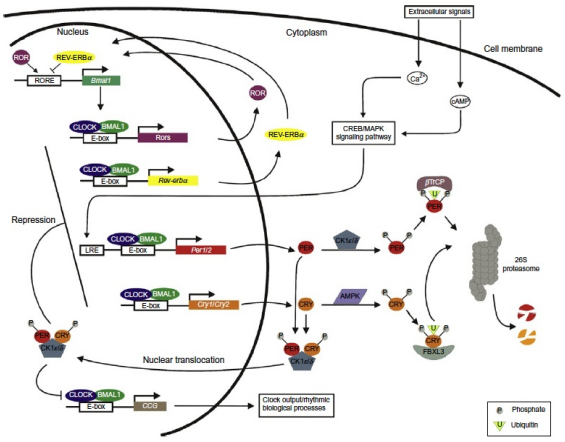

Figure 6. Lowrey and Takahashi's (2011) mechanism diagram of the intracellular circadian oscillator in animals.

One strategy for bringing time into a mechanism diagram is to position different states of the mechanism around a single circle that is marked so as to track time within the 24 hours of one complete cycle. The following diagrams illustrate different ways circadian biologists have instantiated this strategy, each with its own compromises. In Figure 7, Hirano (2013) achieved a visually simple–but conceptually complex–diagram by extracting the CRY cycle from Figure 6 and adding time markers. There are similarities: a single arrow (bottom) suffices to represent the sequence of operations involved in gene expression (transcription of the gene *Cry* in the nucleus, resulting in mRNA that is transported into the cytosol and translated there into CRY); another arrow (top) represents CRY's later translocation into the nucleus. The figures differ in which operations within these spatial compartments are emphasized: Figure 6 shows the dimerization of CRY with PER in the cytosol, whereas Figure 7 incorporates the authors' research on the roles of two other proteins: FBXL21 in stabilizing CRY in the cytosol and FBXL3 in causing the degradation of nuclear CRY. The more important difference for us is the day-night bar at the bottom, which links operations in the cytosol to daytime and those in the nucleus to night. As long as this convention is regarded as simply showing when different operations are at maximum, it captures important timing information. But it is ambiguous regarding the timing of gene expression (which is daytime) and invites false inferences (e.g., that CRY is available in the cytosol only during the day).

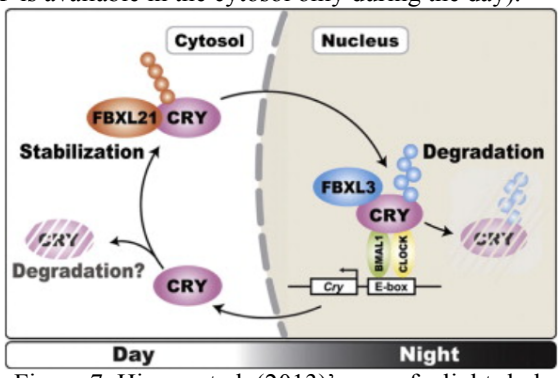

Figure 7. Hirano et al. (2013)'s use of a light-dark bar in a mechanism diagram to indicate time of day when different operations are performed.

The strategy of Figure 7 works only when the focus is on changes involving one component of the full mechanism shown in Figure 6. A related approach that allows for additional parts and operations is to duplicate the mechanism diagram, with appropriate modifications and time markers for each state of interest, and arrange the variations in a circle. In the four panels of Figure 8, for example, Ye, Selby, Ozturk, Annayev, and Sancar (2011) showed the state of key parts of the same mechanism at two times during the

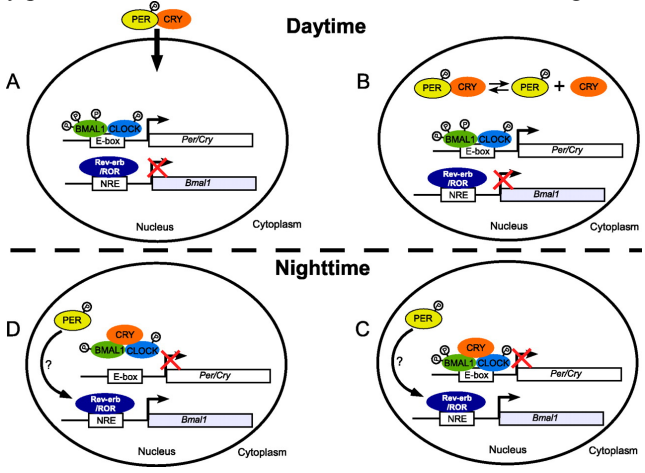

Figure 8. Ye et al.'s (2011) representation of four stages in the daily cycle of the mammalian circadian mechanism.

day and two during the night, focusing especially on the CRY-PER relationship. (Note that panel C is to the right of D, making the arrangement cyclic.) As one moves between panels, the key changes are indicated. For example, as the end of day approaches (moving from panel A to panel B), the PER:CRY dimer enters the nucleus and dissociates into PER and CRY. The red X indicates that *Bmal1* is not being transcribed. With the shift to early night (panel C), CRY binds to BMAL1 and CLOCK on the *Per* and *Cry* E-boxes, stopping their transcription, while *Bmal1* transcription resumes. Finally, these proteins are removed from the E-box (panel D). (For visual simplicity, the conventional rectangular glyphs for the *Per* and *Cry* genes are combined into one.)

One of the challenges in constructing mechanism diagrams is that parts engage in different classes of operations (e.g., movement, promotion of a reaction, phosphorylation) that aren't clearly distinguished by most diagrammatic conventions. For example, the proteins shown in Figure 6 are synthesized in the cytoplasm and transported into the nucleus. On the other hand, the promoter boxes don't themselves move, but rather enable transcription of the downstream gene. Such heterogeneity in the nature of the operations makes it difficult to convey timing information about the different operations in one cohesive way. Some mechanisms, though, can be understood by simply following the transformations of one part. The core mechanism in the circadian clock of cyanobacteria, for example, involves the sequential phosphorylation and dephosphorylation of the protein KaiC at two sites, S431 and T432 (often labeled S and T). The states constitute a cycle that is naturally represented as points around a circle; in Figure 9, for example, a lowercase *p* placed before the S and/or T indicates which sites are phosphorylated in each state.

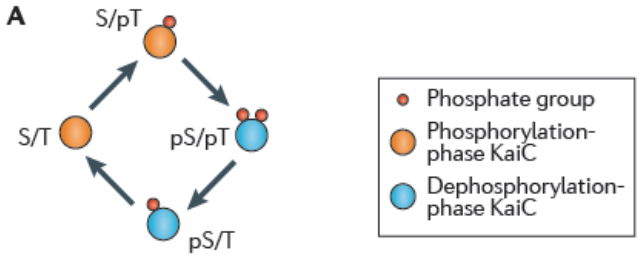

Figure 9. Hogenesch and Ueda's (2011) mechanism diagram of the four stages of KaiC phosphorylation in the cyanobacterial circadian clock.

Although the cycle in Figure 9 is assumed to take 24 hours, the phases of circadian time at which KaiC is in the different states is not indicated. By rotating the image about 135°, Golden, Cassone, and LiWang (2007) were able to show the states around a 24-hour clock face (Figure 10). They also show the roles of two other proteins, KaiA and KaiB. When bound to KaiC, KaiA facilitates the phosphorylation of KaiC whereas KaiB inhibits the activity of KaiA. The shape of the KaiC icon also makes it visually obvious that it is a hexamer, and the faint icons within the clock circle represent the fact that individual monomers can be exchanged between hexamers.

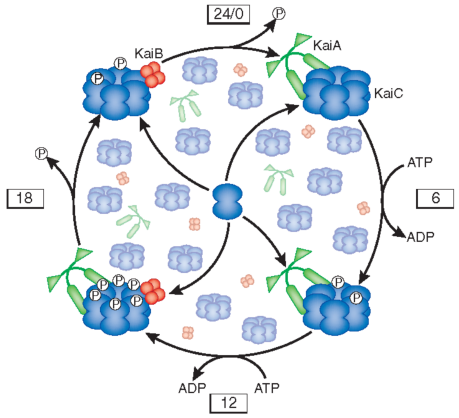

Figure 10. Golden et al.'s (2007) representation of the stages in KaiC phosphorylation showing the circadian times at which KaiC is in each state.

Even though the core mechanism of the cyanobacterial clock consists in the phosphorylation and dephosphorylation of KaiC, this cycle is also thought to be embedded in a transcription-translation feedback cycle (i.e., a negative feedback loop akin to those in Figs. 6-8, but involving KaiA/B/C rather than mammalian clock proteins). Keeping the two Kai cycles distinct but related in a single diagram is challenging, since they involve some of the same parts and operate on the same time-scale. Pattanayek et al. (2011) developed the solution in Figure 11. They show the cycle of phosphorylation and dephosphorylation of Kai C in the upper right ('PTO'), embedded in the larger cycle involving alternation of the chromatin that regulates the transcription and translation of the three Kai genes ('TTFL'). Like other diagrams that stretch the available representational resources, this one carries the risk of inviting false inferences. With no realistic way to mark clock time, for example, it appears that the PTO operates during only one stage of the TTFL. In fact, both of these cycles traverse their sequence of states over a 24-hour period but interact: only when the PTO is in the

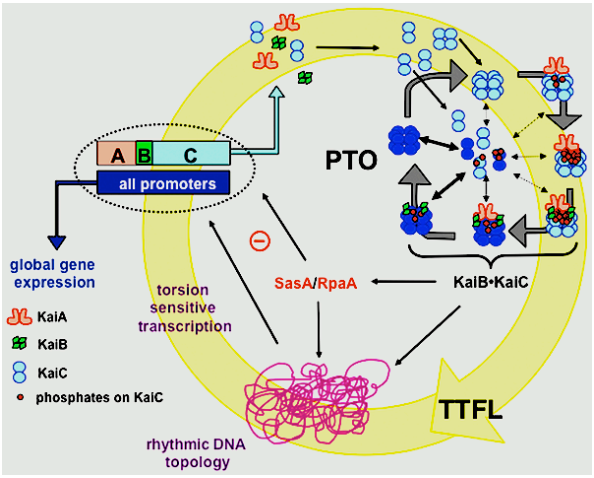

Figure 11. Pattanayek's (2011) mechanism diagram combining the PTO mechanism involving KaiC phosphorylation and a TTFL.

appropriate state does the TTFL proceed to synthesize new Kai proteins. For a reader with the right background knowledge, Figure 11 aptly conveys how the two cycles in the overall mechanism are related to each other.

Conclusion

By examining a variety of diagrams from circadian rhythm research, we have identified a number of ways researchers have solved the problem of representing time. The simplest strategy is to use one spatial dimension to represent time and the other for another variable. Sometimes the two spatial dimensions are both used to represent time, but on different time scales. In that case, other conventions must be adopted to show the amount of activity at each time. Given the importance of the 24-hour cycle for circadian research, representing time in a circle and using a clock face to indicate specific times offers a powerful way to convey information about the phases of different activities. It is more challenging to represent time in mechanism diagrams, which depict the parts, operations, and organization of a mechanism. This is especially true when the mechanism consists of multiple feedback loops. We identified a number of strategies adopted by circadian researchers, such as displaying the states or operations of a mechanism in a circle and adding day/night or time markers.

The diagramming strategies we have identified in scientists' practice pose additional questions for cognitive scientists. Creating as well as consuming the different types of diagrams requires cognitive activities that cognitive science researchers can elucidate. Since each type of diagram involves selectively representing spatial and temporal information, each requires viewers to make appropriate inferences. This can cause novices, and sometimes even experts, to make errors. Studying these errors can help elucidate the reasoning involved. Moreover, diagrams do not function in isolation: different diagrams complement each other's limitations. Learning how scientists produce, understand, misunderstand, and bring together different kinds of diagrams provides indispensable access to scientific reasoning.

Acknowledgments

We gratefully acknowledge the support of National Science Foundation grant 1127640 and numerous contributions by the members of the Center for Chronobiology at UCSD.

References

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421-441.

Bechtel, W., & Richardson, R. C. (1993/2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge, MA: MIT Press. 1993 edition published by Princeton University Press.

Cheng, P. C.-H. (2002). Electrifying diagrams for learning: principles for complex representational systems. *Cognitive Science, 26*, 685-736.

Cheng, P. C.-H. (2011). Probably good diagrams for learning: Representational epistemic recodification of probability theory. *Topics in Cognitive Science, 3*, 475-498.

Ciarleglio, C. M., Gamble, K. L., Axley, J. C., Strauss, B. R., Cohen, J. Y., Colwell, C. S., & McMahon, D. G. (2009). Population encoding by circadian clock neurons organizes circadian behavior. *Journal of Neuroscience, 29*, 1670-1676.

Golden, S. S., Cassone, V. M., & LiWang, A. (2007). Shifting nanoscopic clock gears. *Nature Structural and Molecular Biology, 14*, 362-363.

Gooding, D. C. (2010). Visualizing scientific inference. *Topics in Cognitive Science, 2*, 15-35.

Hardin, P. E., Hall, J. C., & Rosbash, M. (1990). Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature, 343*, 536-540.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Science, 8*, 280-285.

Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in Cognitive Science, 3*, 446-474.

Hirano, A., Yumimoto, K., Tsunematsu, R., Matsumoto, M., Oyama, M., Kozuka-Hata, H., Nakagawa, T., Lanjakornsiripan, D., Nakayama, K. I., & Fukada, Y. (2013). FBXL21 regulates oscillation of the circadian clock through ubiquitination and stabilization of cryptochromes. *Cell, 152*, 1106-1118.

Hogenesch, J. B., & Ueda, H. R. (2011). Understanding systems-level properties: timely stories from the study of clocks. *Nature Reviews Genetics, 12*, 407-416.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1-25.

Nersessian, N. (2008). *Creating scientific concepts*. Cambridge, MA: MIT Press.

Pattanayek, R., Williams, D. R., Rossi, G., Weigand, S., Mori, T., Johnson, C. H., Stewart, P. L., & Egli, M. (2011). Combined SAXS/EM based models of the S. elongatus post-translational circadian oscillator and its interactions with the output His-kinase SasA. *PLoS One, 6*, e23697.

Relógio, A., Westermark, P. O., Wallach, T., Schellenberg, K., Kramer, A., & Herzel, H. (2011). Tuning the mammalian circadian clock: Robust synergy of two loops. *PLoS Computational Biology, 7*, e1002309.

Sheredos, B., Burnston, D., Abrahamsen, A., & Bechtel, W. (2013). Why do biologists use so many diagrams? *Philosophy of Science, 80*, 931-944.

Tversky, B. (2011). Visualizing thought. *Topics in Cognitive Science, 3*, 499-535.

Ueda, H. R. (2007). Systems biology of mammalian circadian clocks. *Cold Spring Harbor Symposia on Quantitative Biology, 72*, 365-380.

Ye, R., Selby, C. P., Ozturk, N., Annayev, Y., & Sancar, A. (2011). Biochemical analysis of the canonical model for the mammalian circadian clock. *The Journal of Biological Chemistry, 286*, 25891-25902.