Classifiers for biosignal interpretation?

-Not beyond toy examples...

T. Teijeiro, P. Félix and J. Presedo

Doctoral Meeting - 23/05/2014

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
CAMPUS VIDA CAMPUS DE EXCELENCIA INTERNACIONAL
CiTiUS Centro Singular de Investigación en Tecnoloxías da Información

Biosignal interpretation: The problem

For decades, a main objective of biosignal processing research has been to provide **classification algorithms** for identifying the underlying physiological phenomena from signal samples by using **pattern recognition**.

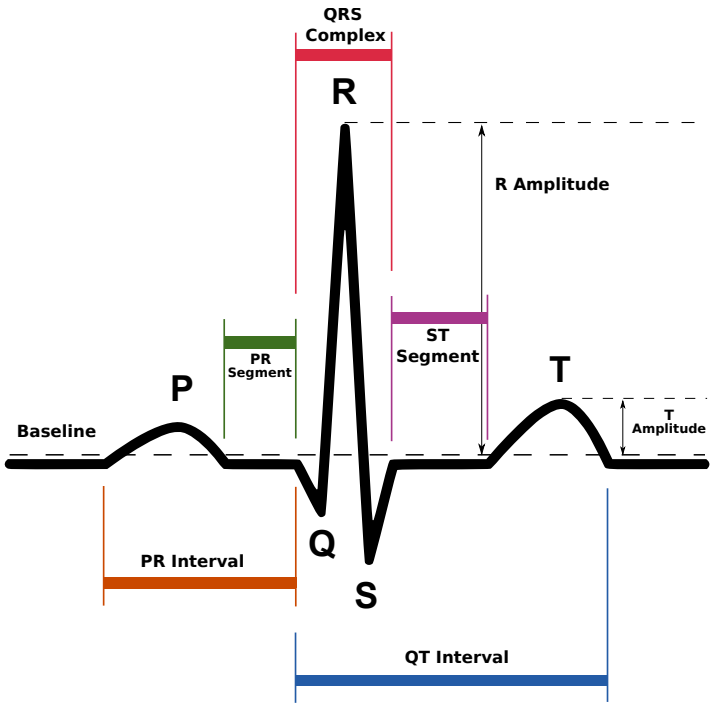
Biosignal interpretation: The problem

For decades, a main objective of biosignal processing research has been to provide **classification algorithms** for identifying the underlying physiological phenomena from signal samples by using **pattern recognition**.

But these signals usually show the **concurrence and interaction between complex processes**, having an extreme intrinsic variability.

Biosignal interpretation: Classical approaches

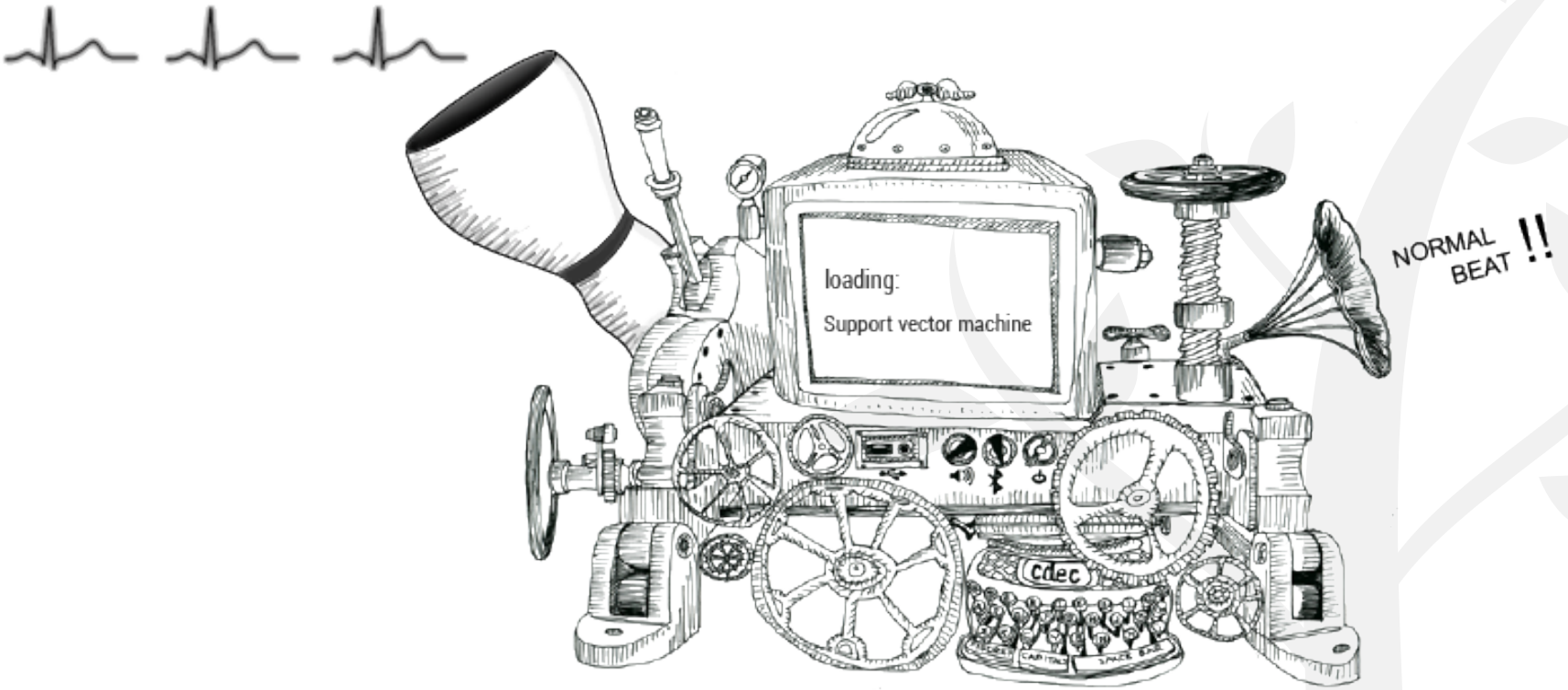
- **Knowledge Based Approaches**: Aim to model the domain knowledge or to simulate an expert reasoning process.



- **IF** {st_segment > 0.1s AND pq_interv < 0.8s} **THEN**...
- **IF** {t_amp / r_amp > 0.5 OR r_amp < 0.4mV} **THEN**...
- **IF** {qt $\in [150, 420]$ AND $\nexists$ P AND QRS < 0.075s} **THEN**...
- ...

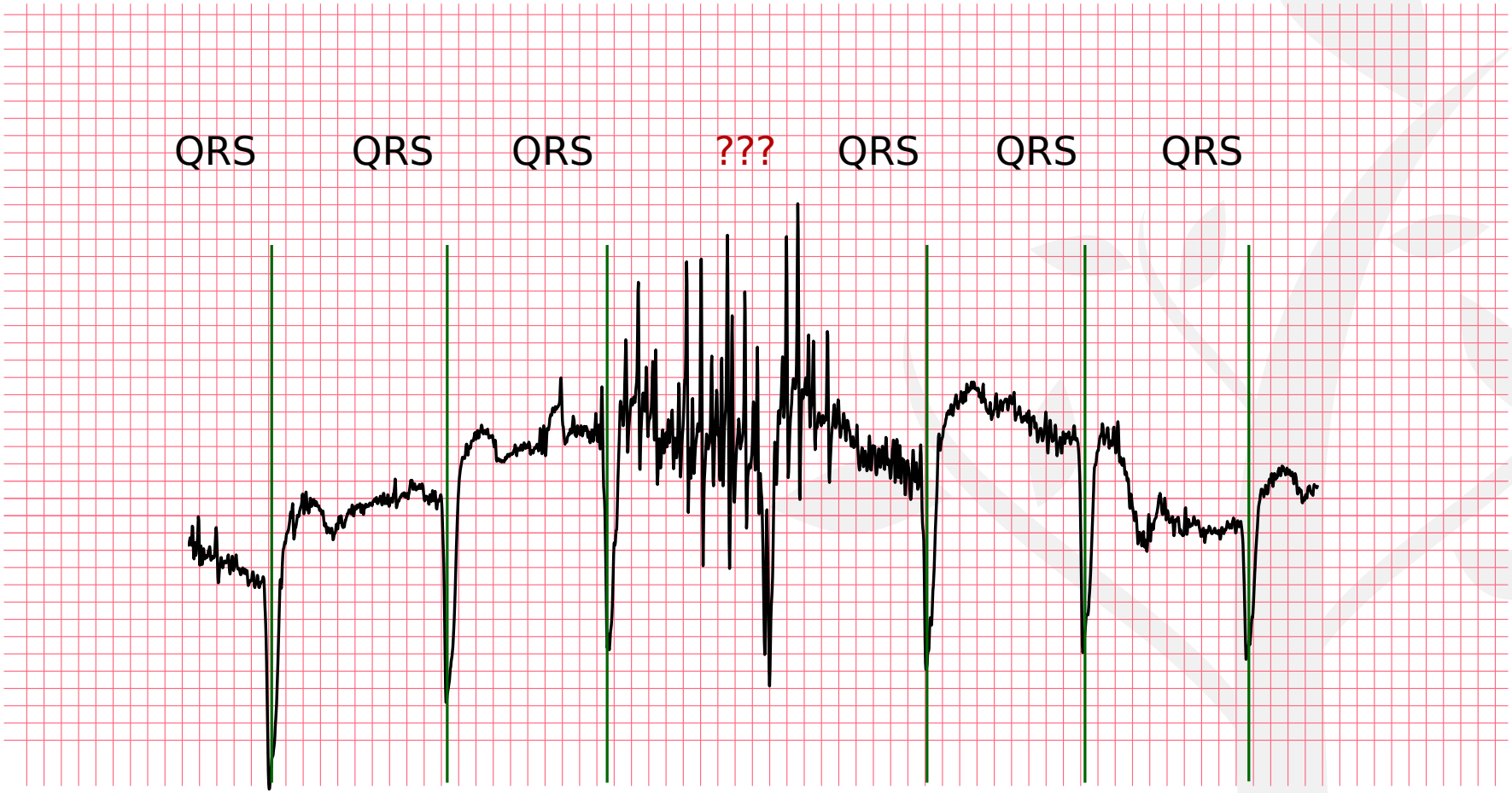
Biosignal interpretation: Classical approaches

- **Learning Based Approaches**: Build a model by estimating the underlying mechanisms that produce the data of a training set.



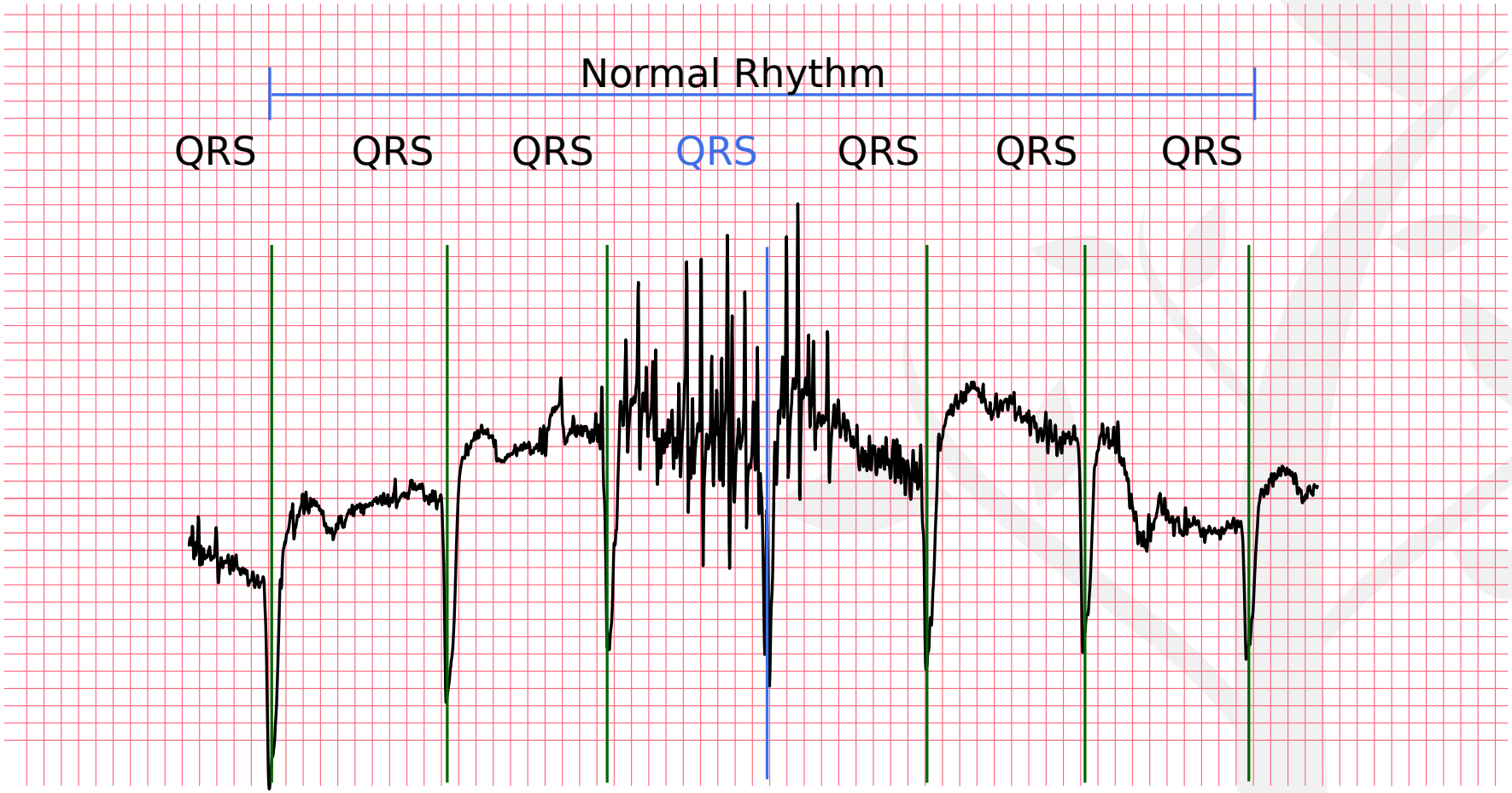
Biosignal interpretation: Classical approaches

Monotonic deductive reasoning cannot fix classification errors with posterior information.



Biosignal interpretation: Classical approaches

Monotonic deductive reasoning cannot fix classification errors with posterior information.



Biosignal interpretation: Classical approaches

Monotonic deductive reasoning cannot fix classification errors with posterior information.

Abductive Reasoning

- Abduction was defined by the philosopher C.S. Peirce as the **process of forming explanatory hypotheses**, illustrated by the following reasoning:

  The surprising fact, C, is observed;
  But if A were true, C would be a matter of course,
  Hence, there is reason to suspect that A is true.

- Which is formalized in the logical formula:

$$\frac{C, A \rightarrow C}{A} \Rightarrow C \mid\!\sim A$$

- But A is not necessarily true. Abduction is **ampliative**, but **fallible.**

Observables

The **observable** $q = \langle \eta, \vec{A}, T_b, T_e \rangle$ is the basic representation entity of our framework, and has the following properties:

- η: Name of the observable.
- $\vec{A} = (A_1, \ldots, A_{n_q})$: Set of attributes to be valued.
- T_b and T_e: Temporal variables representing the beginning and the end of the observable.

Example: QRS observable

$q_{QRS} = \langle \mathtt{QRS}, (\mathtt{shape}, \mathtt{amplitude}), T_b, T_e \rangle$

An observable can be observed in multiple instances called **observations**, defined as $o = \langle \eta, \vec{v}, t_b, t_e \rangle$.

- $o_1 = \langle \mathtt{QRS}, (\mathtt{QS}, 0.94\text{mV}), 00{:}32.123, 00{:}32.201 \rangle$
- $o_2 = \langle \mathtt{QRS}, (\mathtt{rsR'}, 1.92\text{mV}), 00{:}33.054, 00{:}33.139 \rangle$.

Interpretation Problem

Abstraction patterns define an abstraction relation between the observables of a domain, and are generated by the G^{ap} grammars, whose rules are of the following type:

$$H = q_H \rightarrow q[L]C$$
$$C \rightarrow q[L]D \mid q[L] \mid \lambda$$

- H is the hypothesis conjectured by the pattern.
- q is the observable predicted by the rule.
- $[L]$ is a set of temporal constraints between the observable generated by the rule and all the observables previously generated

Given $G \in G^{ap}$, an **abstraction pattern** $P = \langle q_H, M_P, N_P, \Pi_P \rangle$ is a tuple where:

- q_H is the hypothesized observable.
- $M_P = \{q_1, \ldots, q_n\}$ is a set of findings that form the evidence supporting q_H.
- N_P is a temporal network between all the temporal variables of q_H and M_P.
- Π_P is an observation procedure to compute the attribute values of q_H.

$$\forall q_i \in M_P : q_i \mid\!\prec q_H$$

The knowledge base of a particular domain is represented in a **domain model** $\mathcal{M} = \langle \mathcal{Q}, \mathcal{P} \rangle$ where:

- $\mathcal{Q}$ Is a set of domain observables.
- $\mathcal{P}$ Is a set of abstraction patterns involving those observables.

An **interpretation problem** $IP = \langle \mathcal{O}, \mathcal{M} \rangle$ is defined by a set of initial observations $\mathcal{O}$ and a domain model $\mathcal{M}$.

Interpretation Problem

Abstraction patterns define an abstraction relation between the observables of a domain, and are generated by the G^{ap} grammars, whose rules are of the following type:

$$H = q_H \rightarrow q[L]C$$
$$C \rightarrow q[L]D \mid q[L] \mid \lambda$$

- H is the hypothesis conjectured by the pattern.
- q is the observable predicted by the rule.
- $[L]$ is a set of temporal constraints between the observable generated by the rule and all the observables previously generated

Given $G \in G^{ap}$, an **abstraction pattern** $P = \langle q_H, M_P, N_P, \Pi_P \rangle$ is a tuple where:

- q_H is the hypothesized observable.
- $M_P = \{q_1, \ldots, q_n\}$ is a set of findings that form the evidence supporting q_H.
- N_P is a temporal network between all the temporal variables of q_H and M_P.
- Π_P is an observation procedure to compute the attribute values of q_H.

$$\forall q_i \in M_P : q_i \Subset q_H$$

The knowledge base of a particular domain is represented in a **domain model** $\mathcal{M} = \langle \mathcal{Q}, \mathcal{P} \rangle$ where:

- $\mathcal{Q}$ Is a set of domain observables.
- $\mathcal{P}$ Is a set of abstraction patterns involving those observables.

An **interpretation problem** $IP = \langle \mathcal{O}, \mathcal{M} \rangle$ is defined by a set of initial observations $\mathcal{O}$ and a domain model $\mathcal{M}$.

Interpretation Problem

Abstraction patterns define an abstraction relation between the observables of a domain, and are generated by the G^{ap} grammars, whose rules are of the following type:

$$H = q_H \rightarrow q[L]C$$
$$C \rightarrow q[L]D \mid q[L] \mid \lambda$$

- H is the hypothesis conjectured by the pattern.
- q is the observable predicted by the rule.
- $[L]$ is a set of temporal constraints between the observable generated by the rule and all the observables previously generated

Given $G \in G^{ap}$, an **abstraction pattern** $P = \langle q_H, M_P, N_P, \Pi_P \rangle$ is a tuple where:

- q_H is the hypothesized observable.
- $M_P = \{q_1, \ldots, q_n\}$ is a set of findings that form the evidence supporting q_H.
- N_P is a temporal network between all the temporal variables of q_H and M_P.
- Π_P is an observation procedure to compute the attribute values of q_H.

$$\forall q_i \in M_P : q_i \Lsh q_H$$

The knowledge base of a particular domain is represented in a **domain model** $\mathcal{M} = \langle \mathcal{Q}, \mathcal{P} \rangle$ where:

- $\mathcal{Q}$ Is a set of domain observables.
- $\mathcal{P}$ Is a set of abstraction patterns involving those observables.

An **interpretation problem** $IP = \langle \mathcal{O}, \mathcal{M} \rangle$ is defined by a set of initial observations $\mathcal{O}$ and a domain model $\mathcal{M}$.

Interpretation Problem

Abstraction patterns define an abstraction relation between the observables of a domain, and are generated by the G^{ap} grammars, whose rules are of the following type:

$$H = q_H \rightarrow q[L]C$$
$$C \rightarrow q[L]D \mid q[L] \mid \lambda$$

- H is the hypothesis conjectured by the pattern.
- q is the observable predicted by the rule.
- $[L]$ is a set of temporal constraints between the observable generated by the rule and all the observables previously generated

Given $G \in G^{ap}$, an **abstraction pattern** $P = \langle q_H, M_P, N_P, \Pi_P \rangle$ is a tuple where:

- q_H is the hypothesized observable.
- $M_P = \{q_1, \ldots, q_n\}$ is a set of findings that form the evidence supporting q_H.
- N_P is a temporal network between all the temporal variables of q_H and M_P.
- Π_P is an observation procedure to compute the attribute values of q_H.

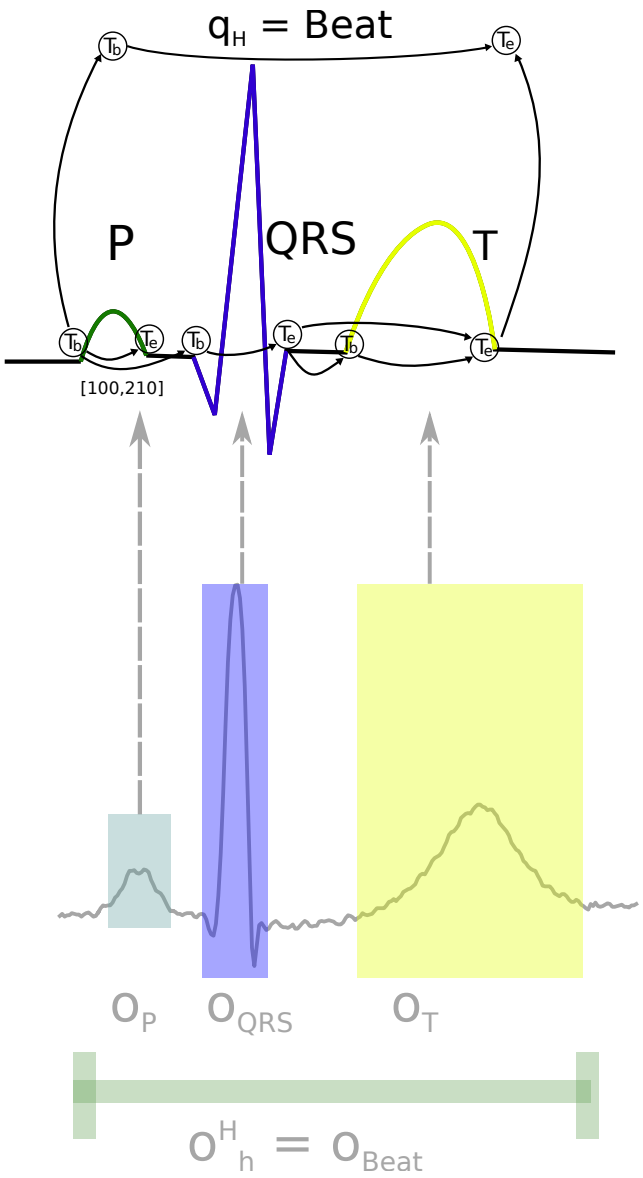
$$\forall q_i \in M_P : q_i \Lsh q_H$$

The knowledge base of a particular domain is represented in a **domain model** $\mathcal{M} = \langle \mathcal{Q}, \mathcal{P} \rangle$ where:

- $\mathcal{Q}$ Is a set of domain observables.
- $\mathcal{P}$ Is a set of abstraction patterns involving those observables.

An **interpretation problem** $IP = \langle \mathcal{O}, \mathcal{M} \rangle$ is defined by a set of initial observations $\mathcal{O}$ and a domain model $\mathcal{M}$.

Hypotheses and Interpretations

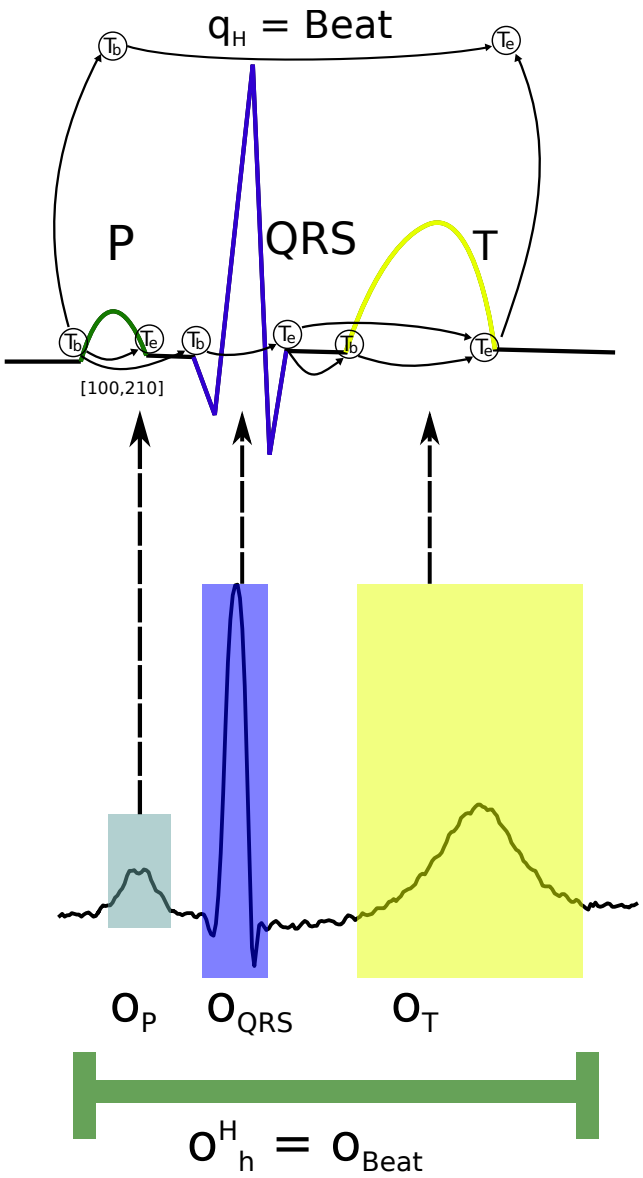


An **abstraction hypothesis** is a tuple $h = \langle o_h^H, P_h, O_h, \leftarrow_h, N_h \rangle$, where:

- $o_h^H = \Pi_{P_h}(O_h)$ is the hypothesized observation.
- P_h is the abstraction pattern generating the hypothesis.
- $O_h = \{o_1, \ldots, o_n\}$ is the set of observations abstracted by the hypothesis.
- $\leftarrow_h$ is an abstraction relation between the M_{P_h} findings and the abstracted observations O_h.
- N_h Is a temporal network between the temporal variables of the hypothesis and the evidence.

An **interpretation** $I = \{h_1, \ldots, h_m\}$ is a set of consistent abstraction hypotheses.

Hypotheses and Interpretations

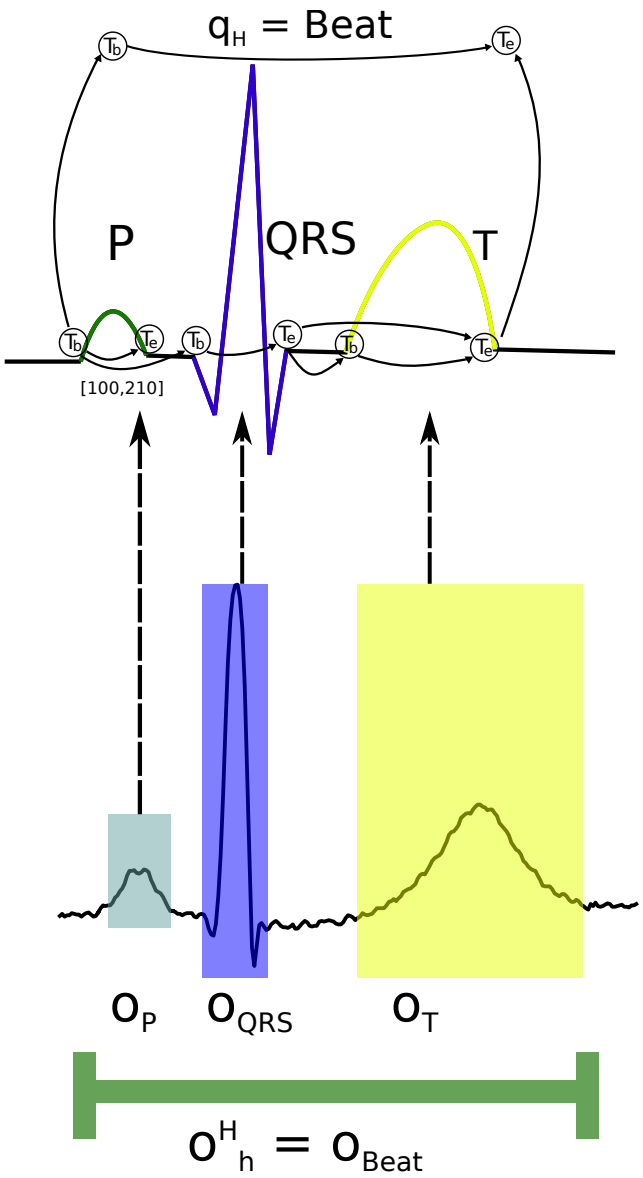


An **abstraction hypothesis** is a tuple $h = \langle o_h^H, P_h, O_h, \leftarrow_h, N_h \rangle$, where:

- $o_h^H = \Pi_{P_h}(O_h)$ is the hypothesized observation.
- P_h is the abstraction pattern generating the hypothesis.
- $O_h = \{o_1, \ldots, o_n\}$ is the set of observations abstracted by the hypothesis.
- $\leftarrow_h$ is an abstraction relation between the M_{P_h} findings and the abstracted observations O_h.
- N_h Is a temporal network between the temporal variables of the hypothesis and the evidence.

An **interpretation** $I = \{h_1, \ldots, h_m\}$ is a set of consistent abstraction hypotheses.

Hypotheses and Interpretations



An **abstraction hypothesis** is a tuple $h = \langle o_h^H, P_h, O_h, \leftarrow_h, N_h \rangle$, where:

- $o_h^H = \Pi_{P_h}(O_h)$ is the hypothesized observation.
- P_h is the abstraction pattern generating the hypothesis.
- $O_h = \{o_1, \ldots, o_n\}$ is the set of observations abstracted by the hypothesis.
- $\leftarrow_h$ is an abstraction relation between the M_{P_h} findings and the abstracted observations O_h.
- N_h Is a temporal network between the temporal variables of the hypothesis and the evidence.

An **interpretation** $I = \{h_1, \ldots, h_m\}$ is a set of consistent abstraction hypotheses.

Interpretation algorithm

Theorem:

Finding the solution of an interpretation problem is NP-Hard.

In order to reduce the complexity of the problem, we pose the interpretation algorithm as an *heuristic search process*, guided by the following general principles:

1. **Coverage principle**: An interpretation explaining more initial observations is better.
2. **Simplicity principle**: An interpretation with fewer hypotheses is better (Occam's razor).
3. **Abstraction principle**: An interpretation that uses terms of higher abstraction levels is better.
4. **Predictive principle**: An interpretation that predicts future observations is better.

Interpretation algorithm

Theorem:

Finding the solution of an interpretation problem is NP-Hard.

In order to reduce the complexity of the problem, we pose the interpretation algorithm as an *heuristic search process*, guided by the following general principles:

1. **Coverage principle**: An interpretation explaining more initial observations is better.
2. **Simplicity principle**: An interpretation with fewer hypotheses is better (Occam's razor).
3. **Abstraction principle**: An interpretation that uses terms of higher abstraction levels is better.
4. **Predictive principle**: An interpretation that predicts future observations is better.

Interpretation algorithm

Theorem:

Finding the solution of an interpretation problem is NP-Hard.

In order to reduce the complexity of the problem, we pose the interpretation algorithm as an *heuristic search process*, guided by the following general principles:

1. **Coverage principle**: An interpretation explaining more initial observations is better.
2. **Simplicity principle**: An interpretation with fewer hypotheses is better (Occam's razor).
3. **Abstraction principle**: An interpretation that uses terms of higher abstraction levels is better.
4. **Predictive principle**: An interpretation that predicts future observations is better.

Interpretation algorithm

Theorem:

Finding the solution of an interpretation problem is NP-Hard.

In order to reduce the complexity of the problem, we pose the interpretation algorithm as an *heuristic search process*, guided by the following general principles:

1. **Coverage principle**: An interpretation explaining more initial observations is better.
2. **Simplicity principle**: An interpretation with fewer hypotheses is better (Occam's razor).
3. **Abstraction principle**: An interpretation that uses terms of higher abstraction levels is better.
4. **Predictive principle**: An interpretation that predicts future observations is better.

Interpretation algorithm

Theorem:

Finding the solution of an interpretation problem is NP-Hard.

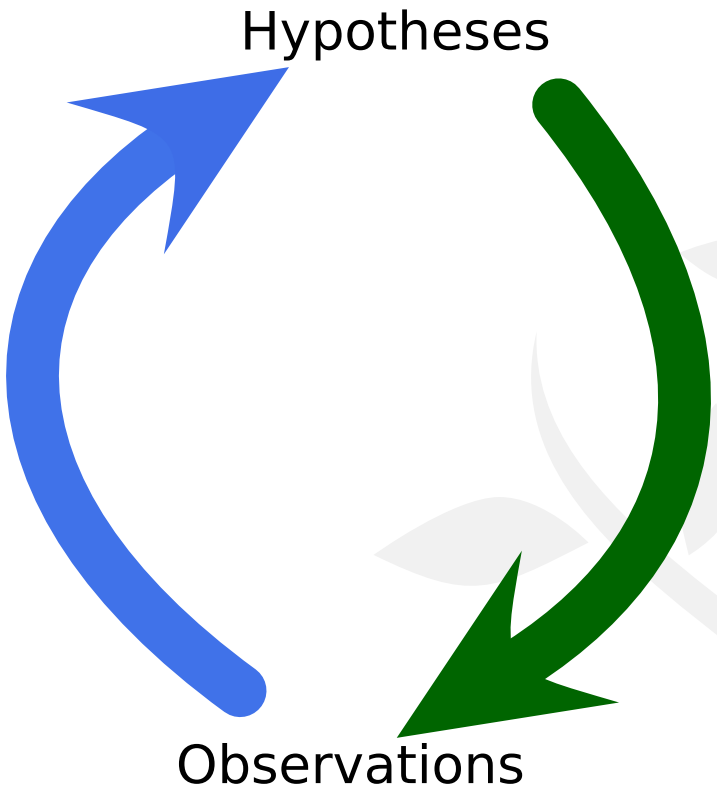
In order to reduce the complexity of the problem, we pose the interpretation algorithm as an *heuristic search process*, guided by the following general principles:

1. **Coverage principle**: An interpretation explaining more initial observations is better.
2. **Simplicity principle**: An interpretation with fewer hypotheses is better (Occam's razor).
3. **Abstraction principle**: An interpretation that uses terms of higher abstraction levels is better.
4. **Predictive principle**: An interpretation that predicts future observations is better.

Building an interpretation

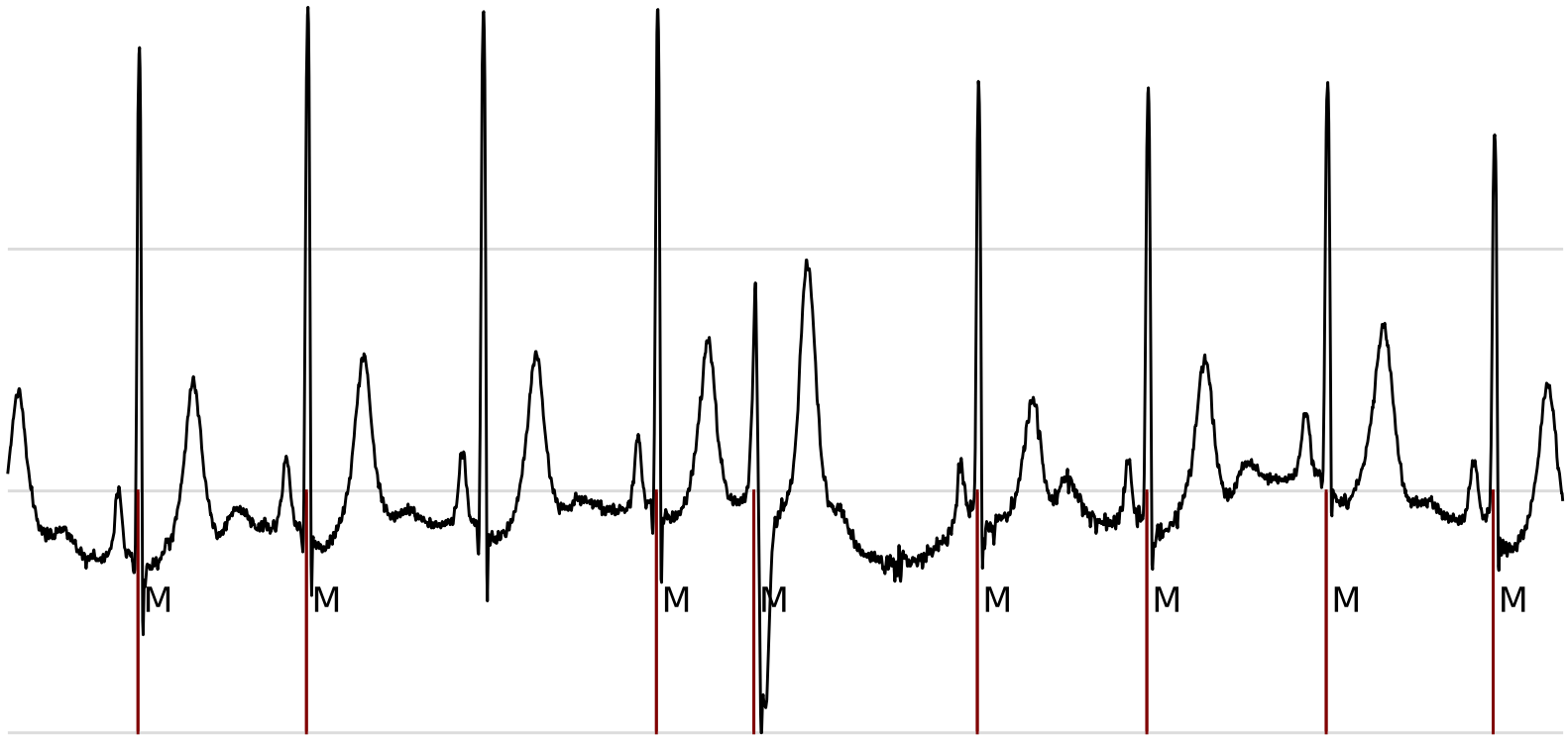
The interpretation is posed as a hypothesis-and-test cycle, in which a **focus of attention** determines what is the next reasoning step:

- **Abduction:** Conjectures an hypothesis explaining the observation that is in the focus of attention.
- **Subsumption:** Matches an existing observation with a finding of an existing hypothesis.

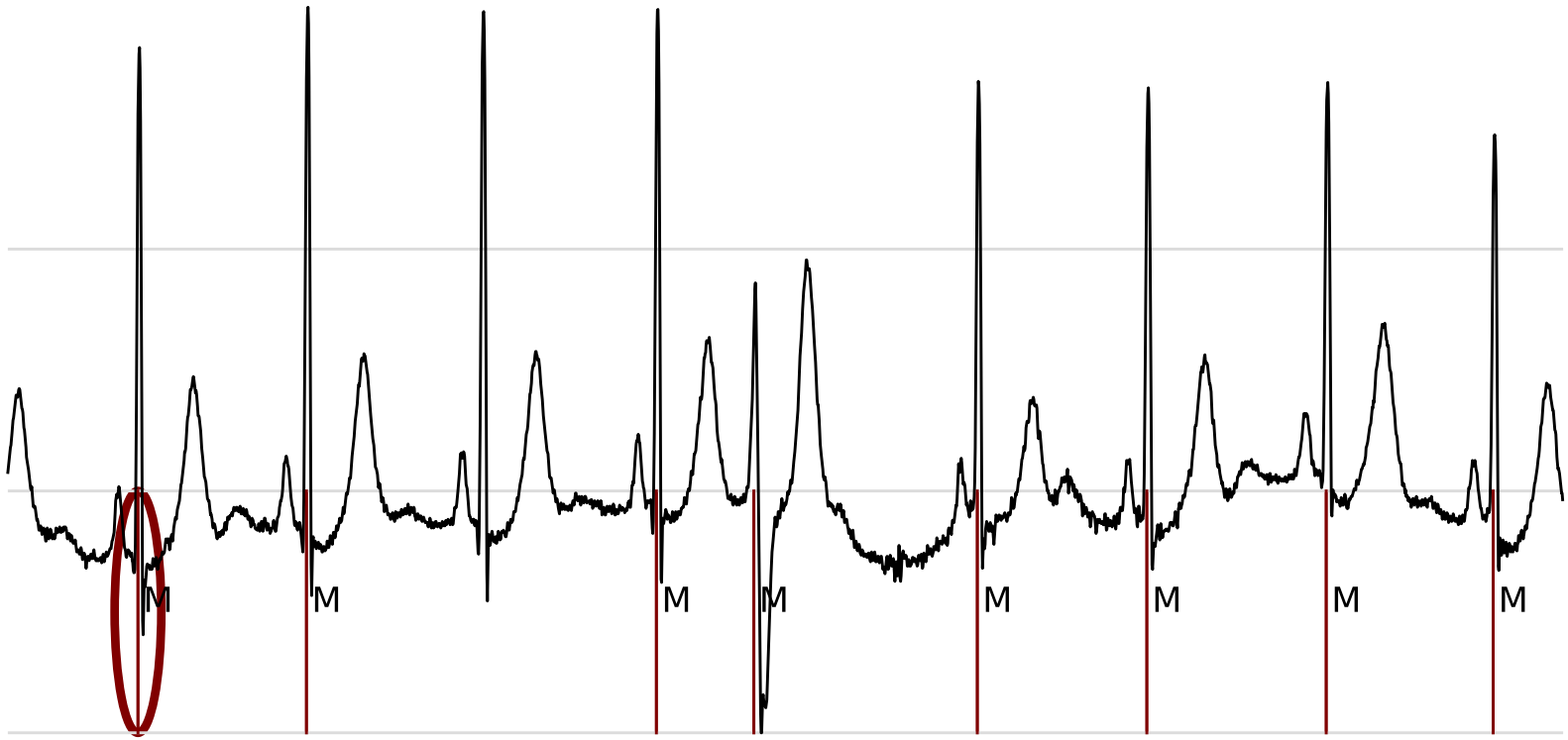


- **Deduction:** Looks for an observation predicted by the abstraction pattern that conjectured a hypothesis.

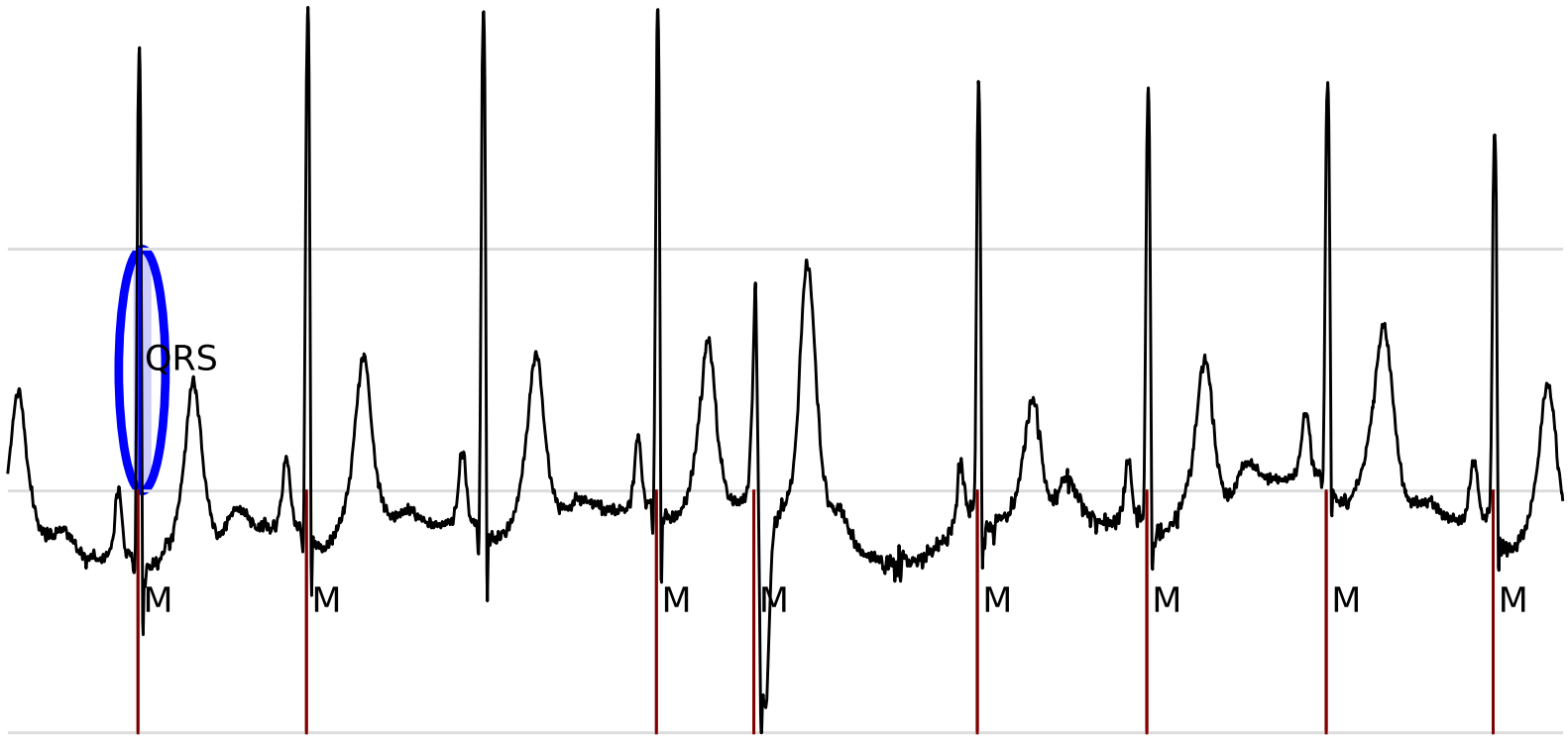
Interpretation example



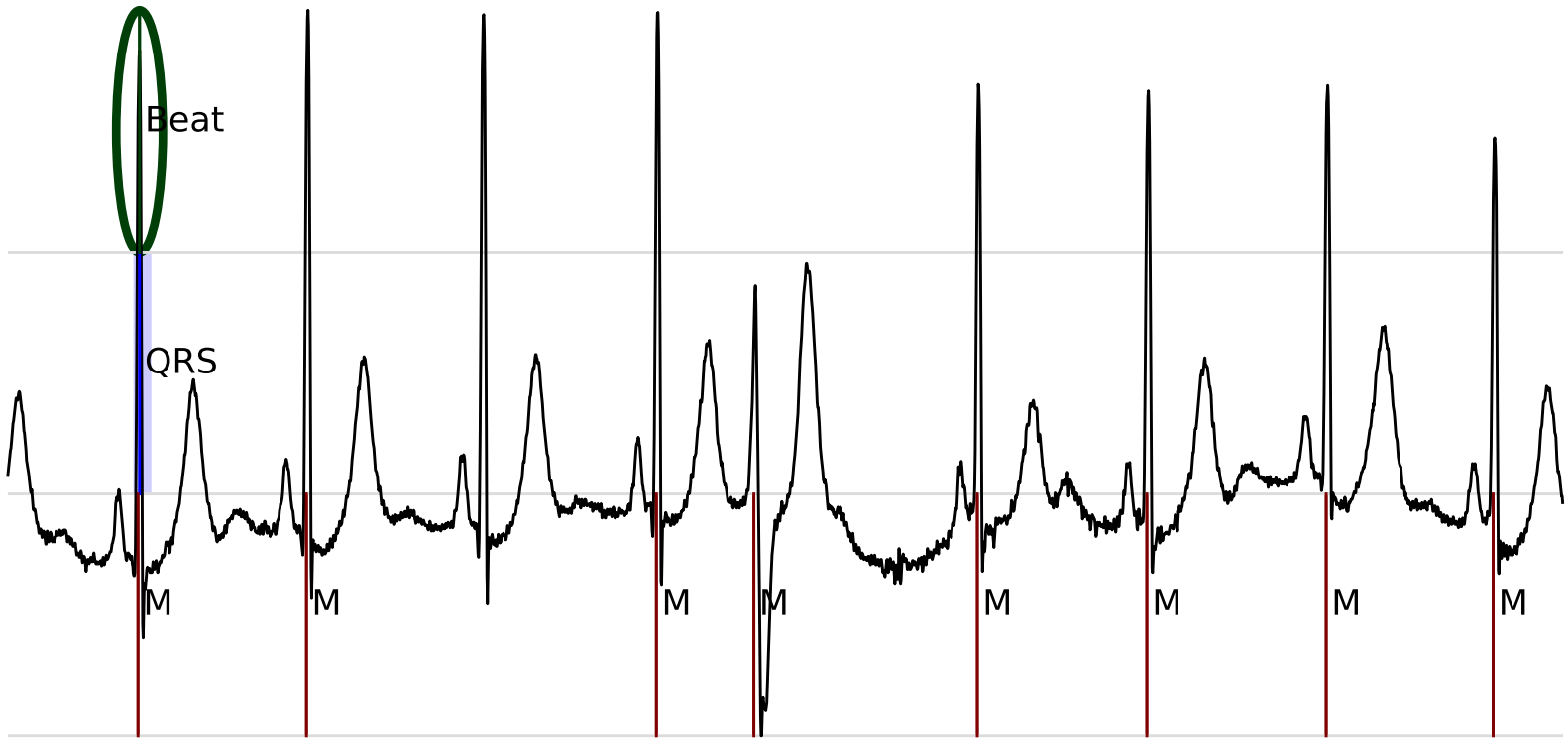
Interpretation example



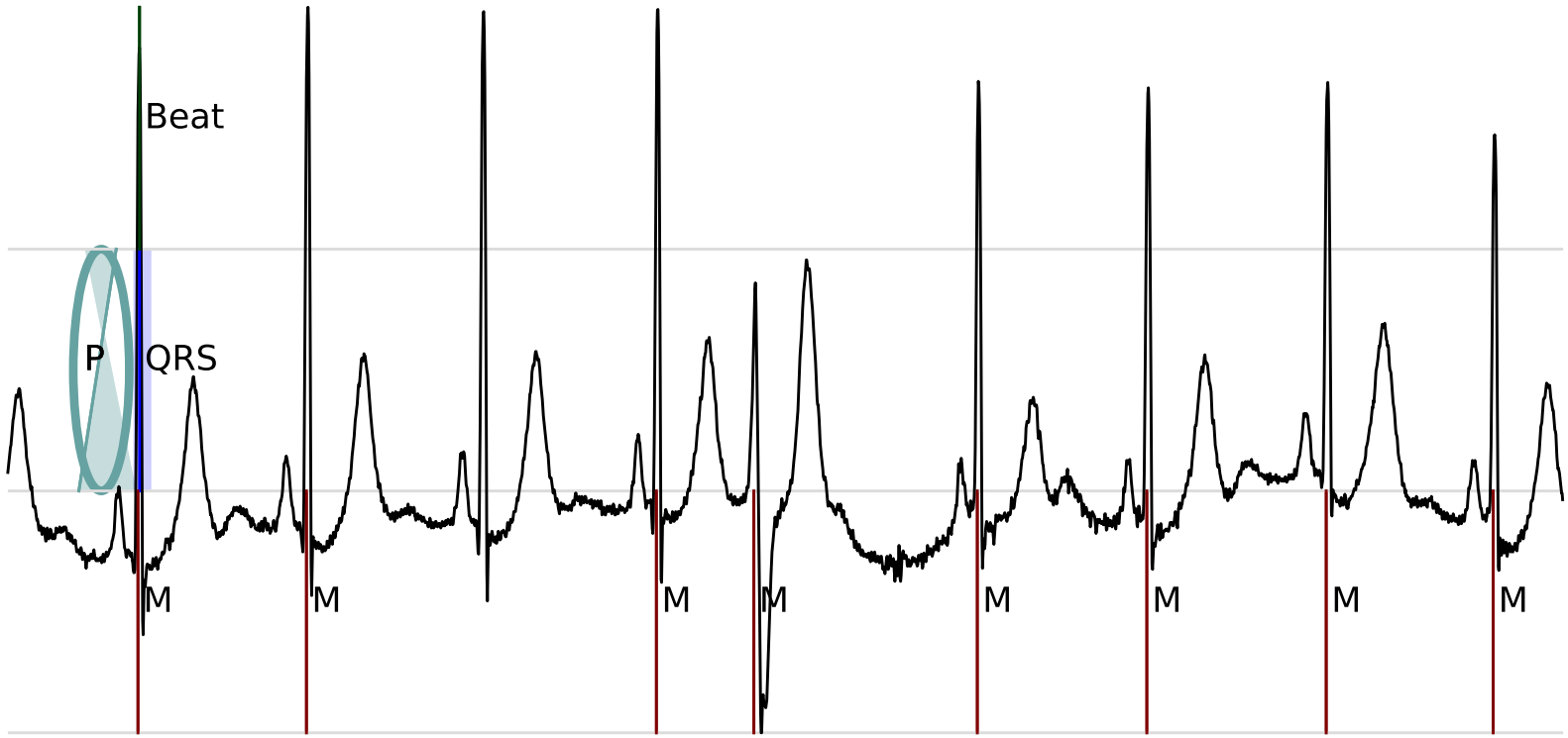
Interpretation example



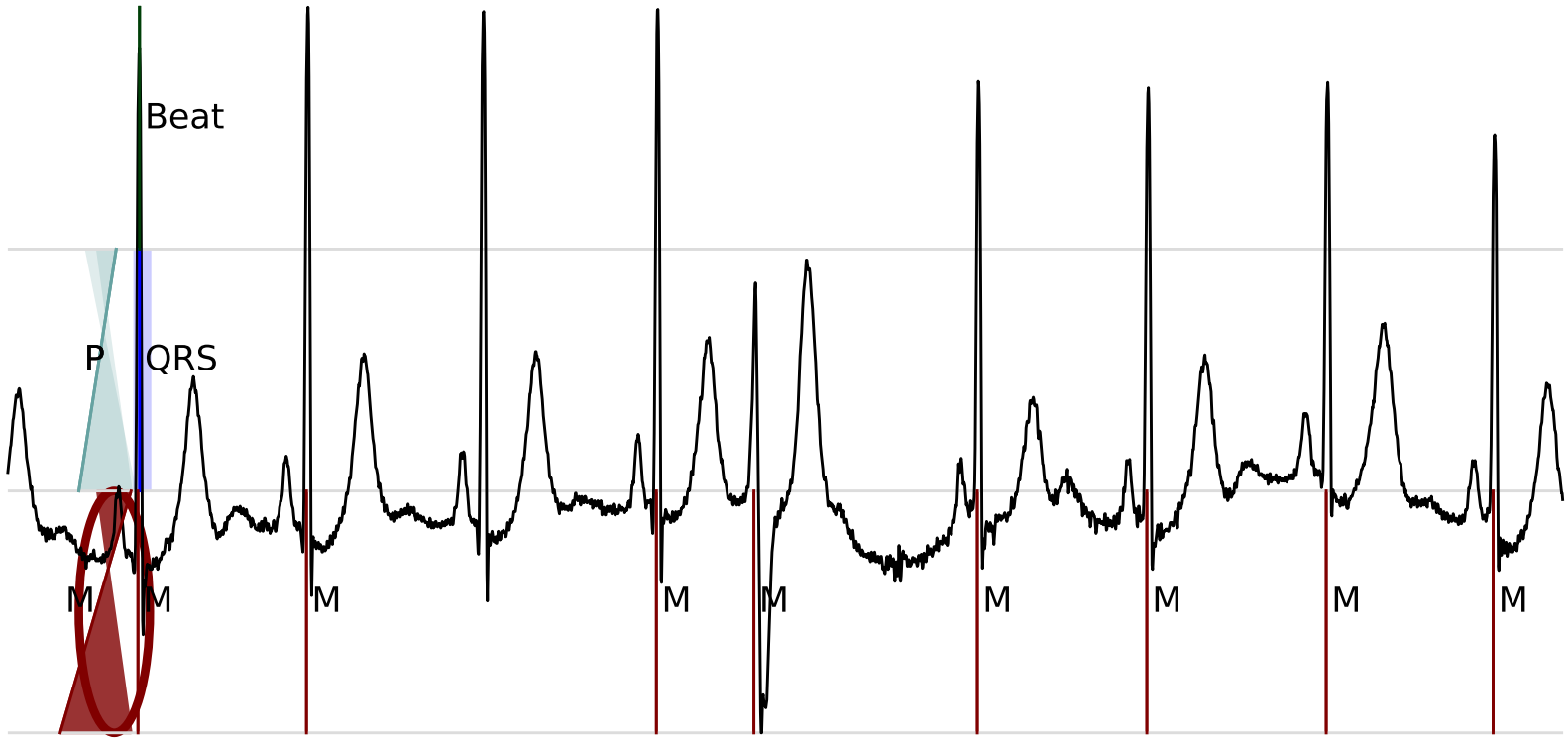
Interpretation example



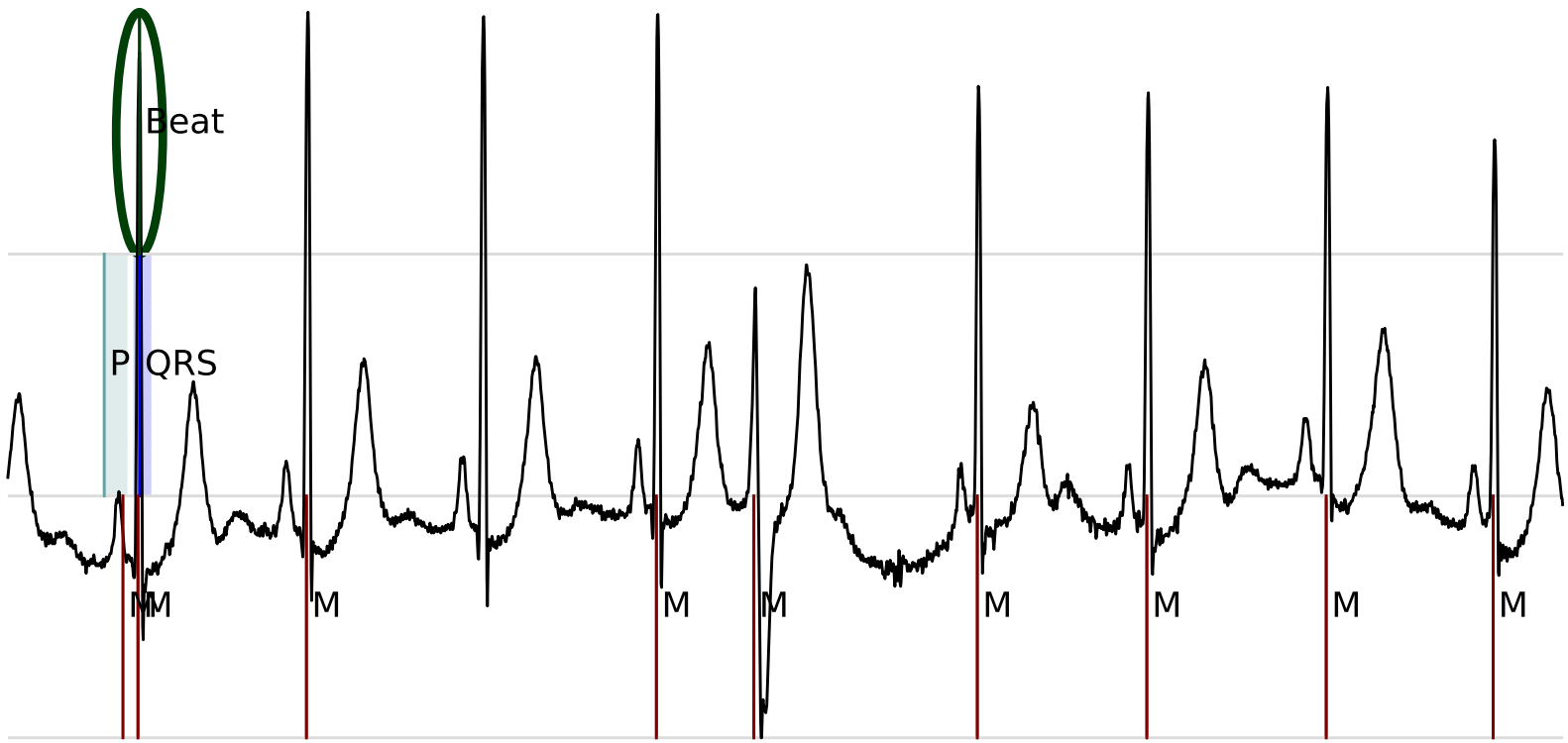
Interpretation example



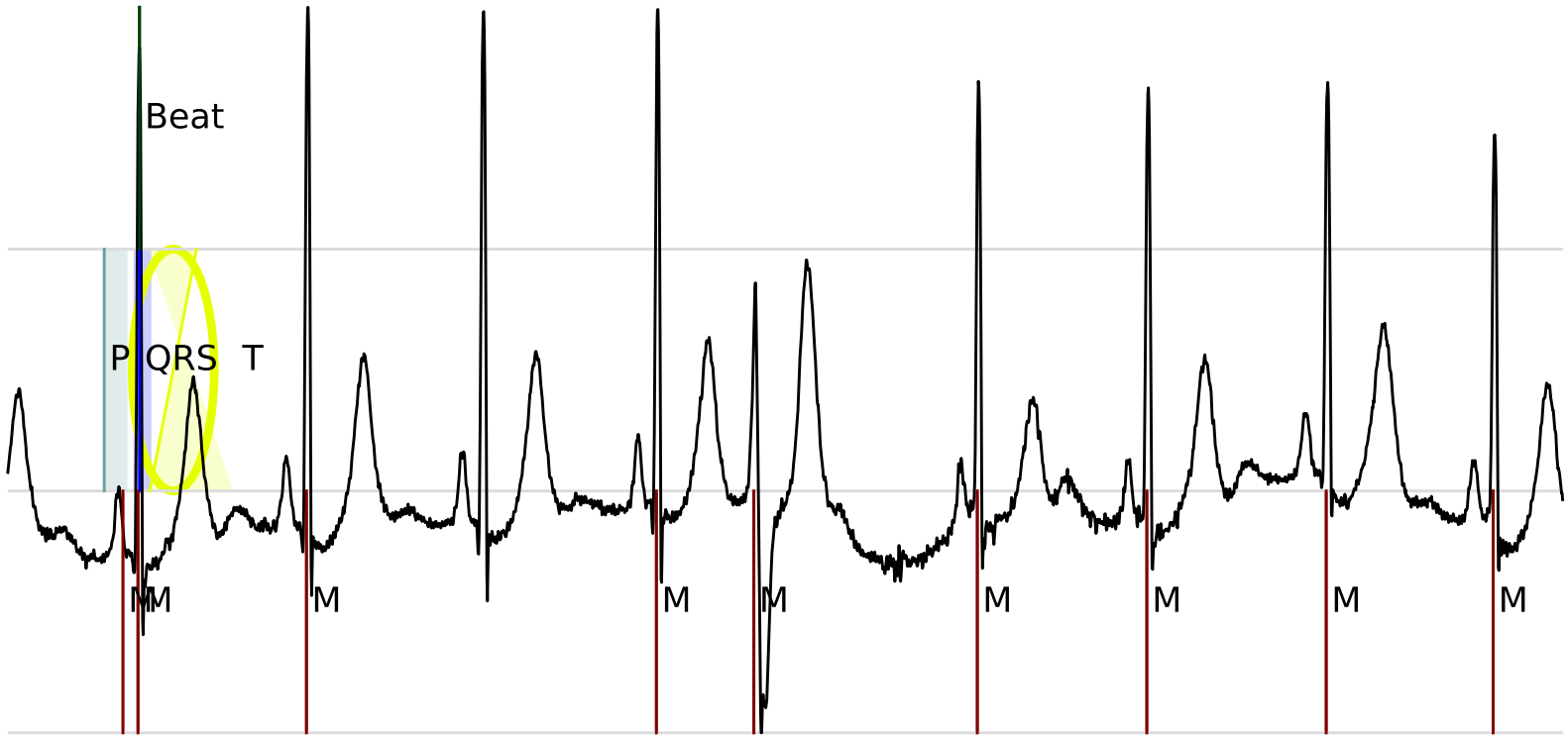
Interpretation example



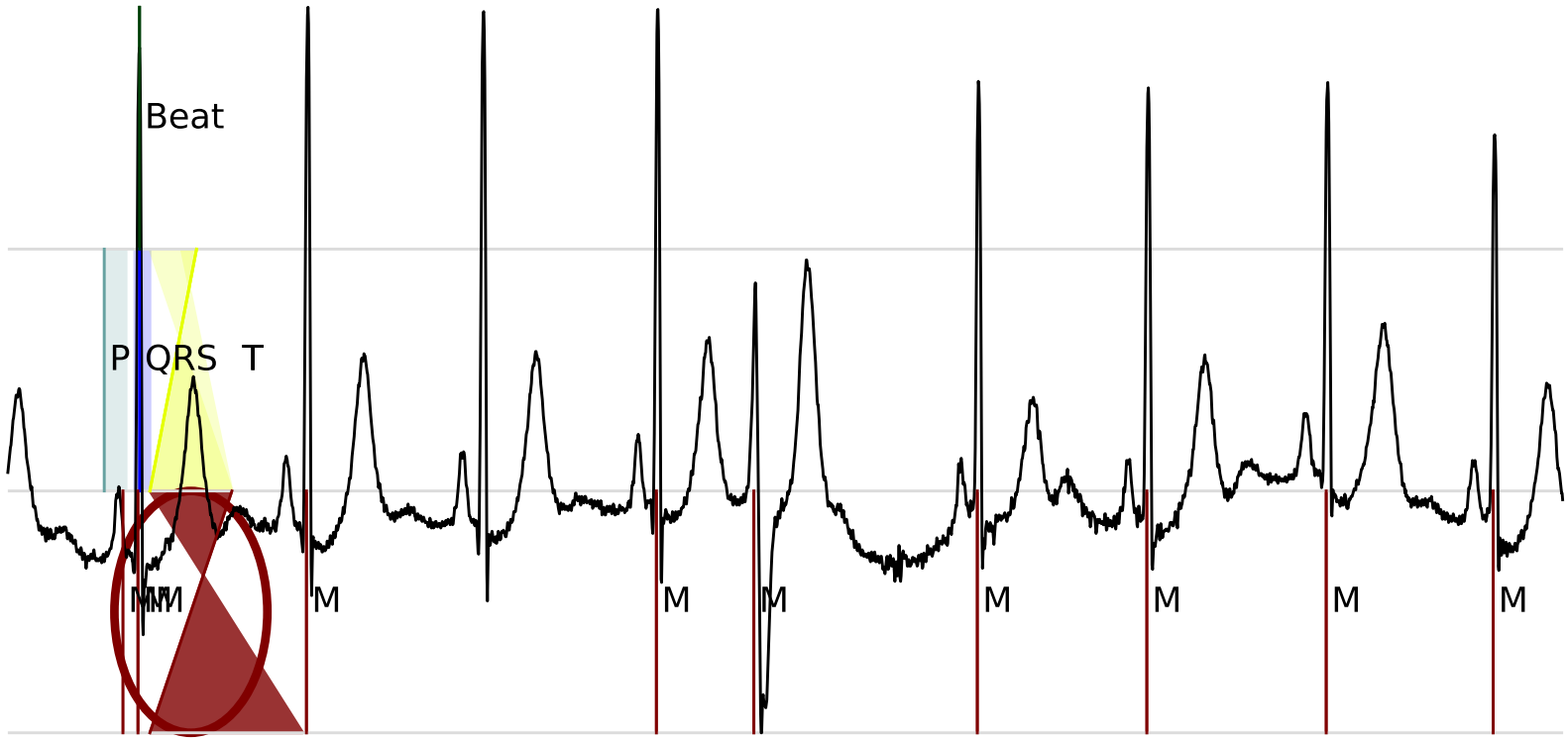
Interpretation example



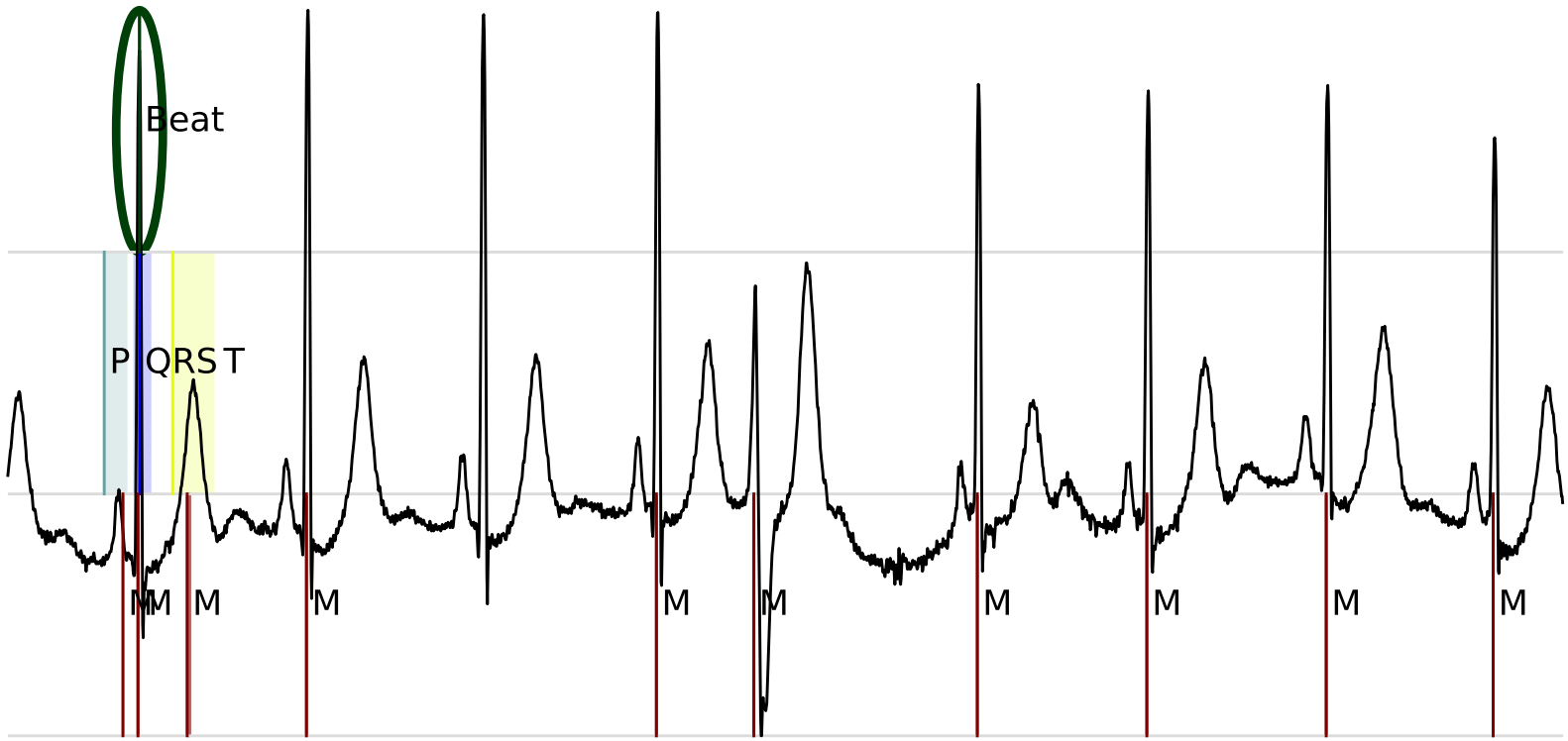
Interpretation example



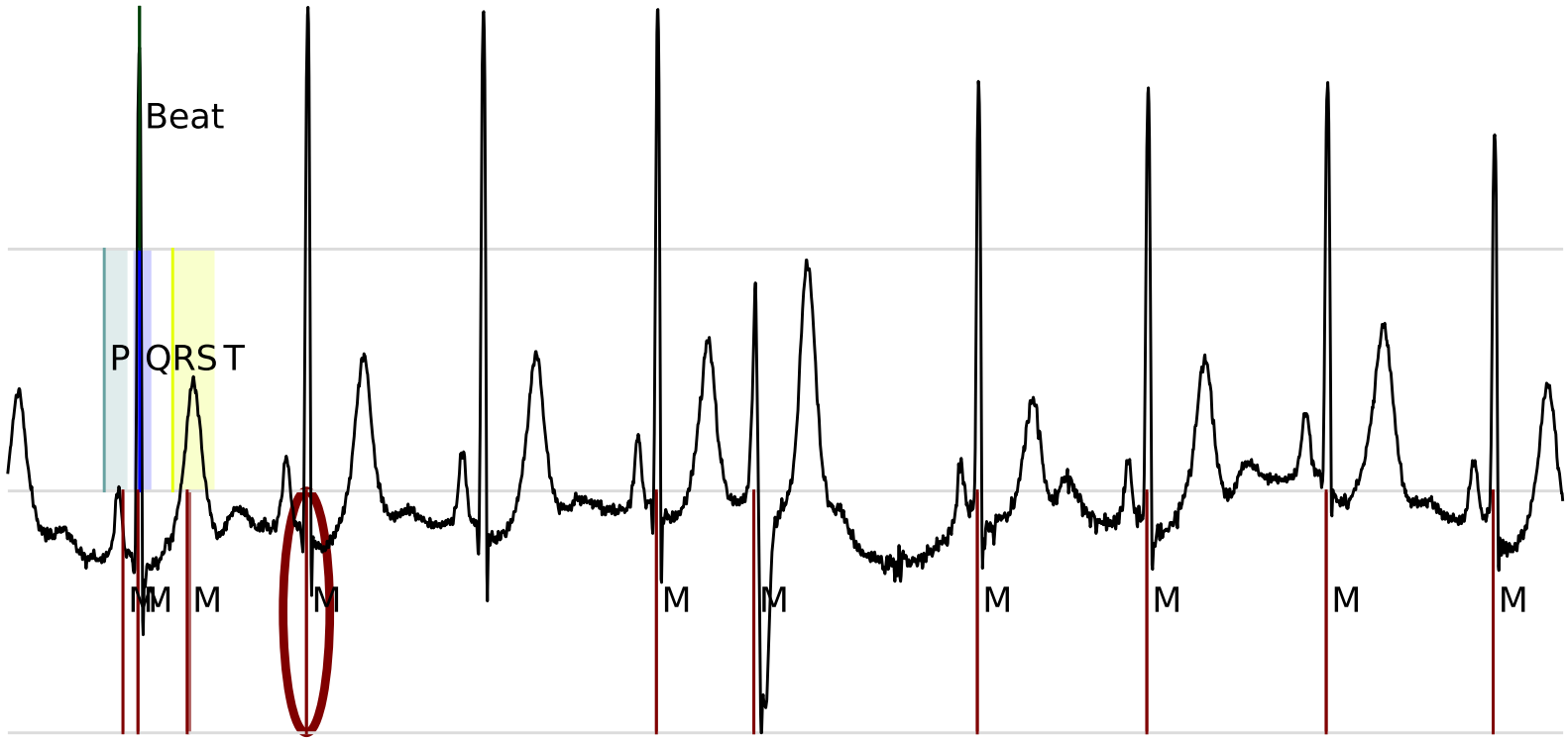
Interpretation example



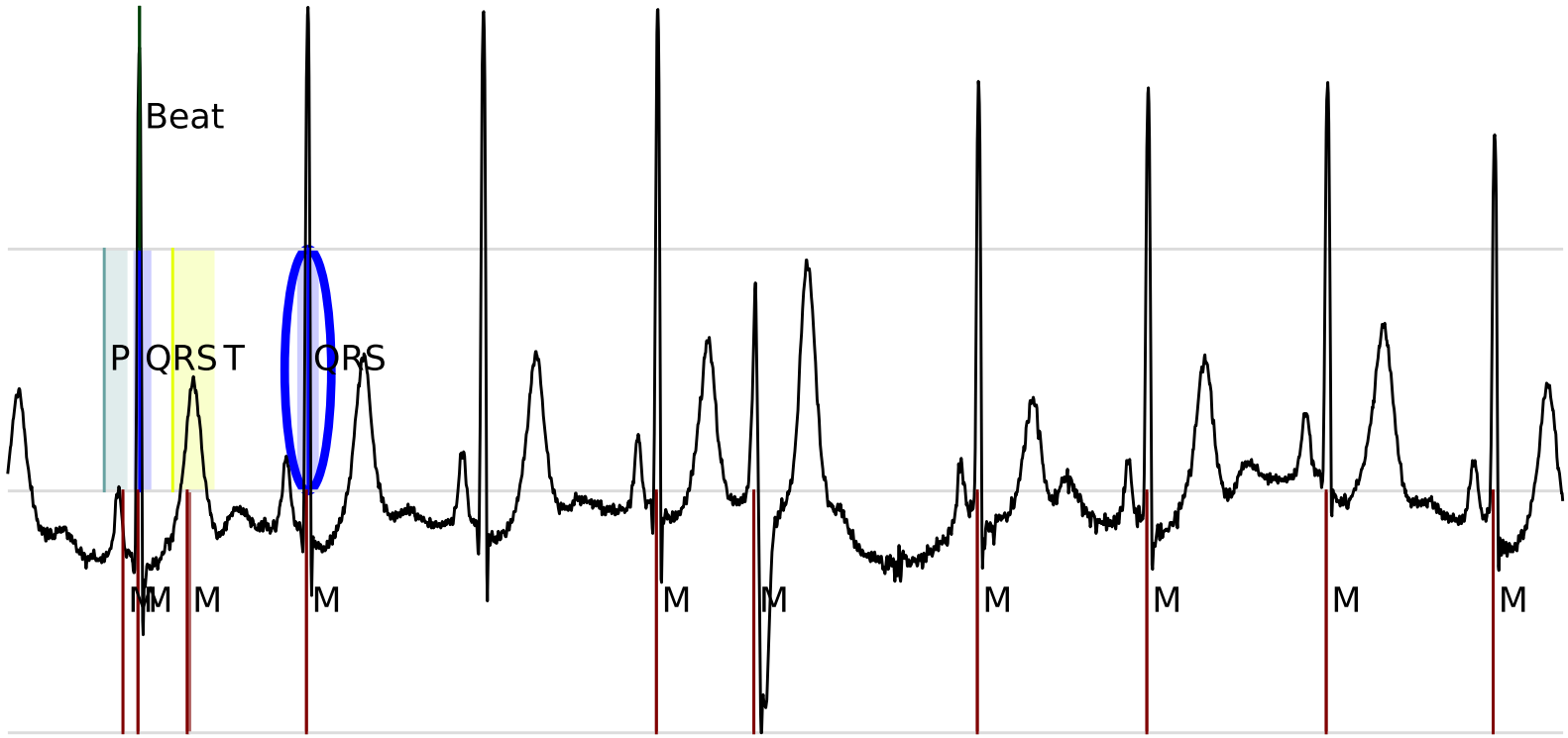
Interpretation example



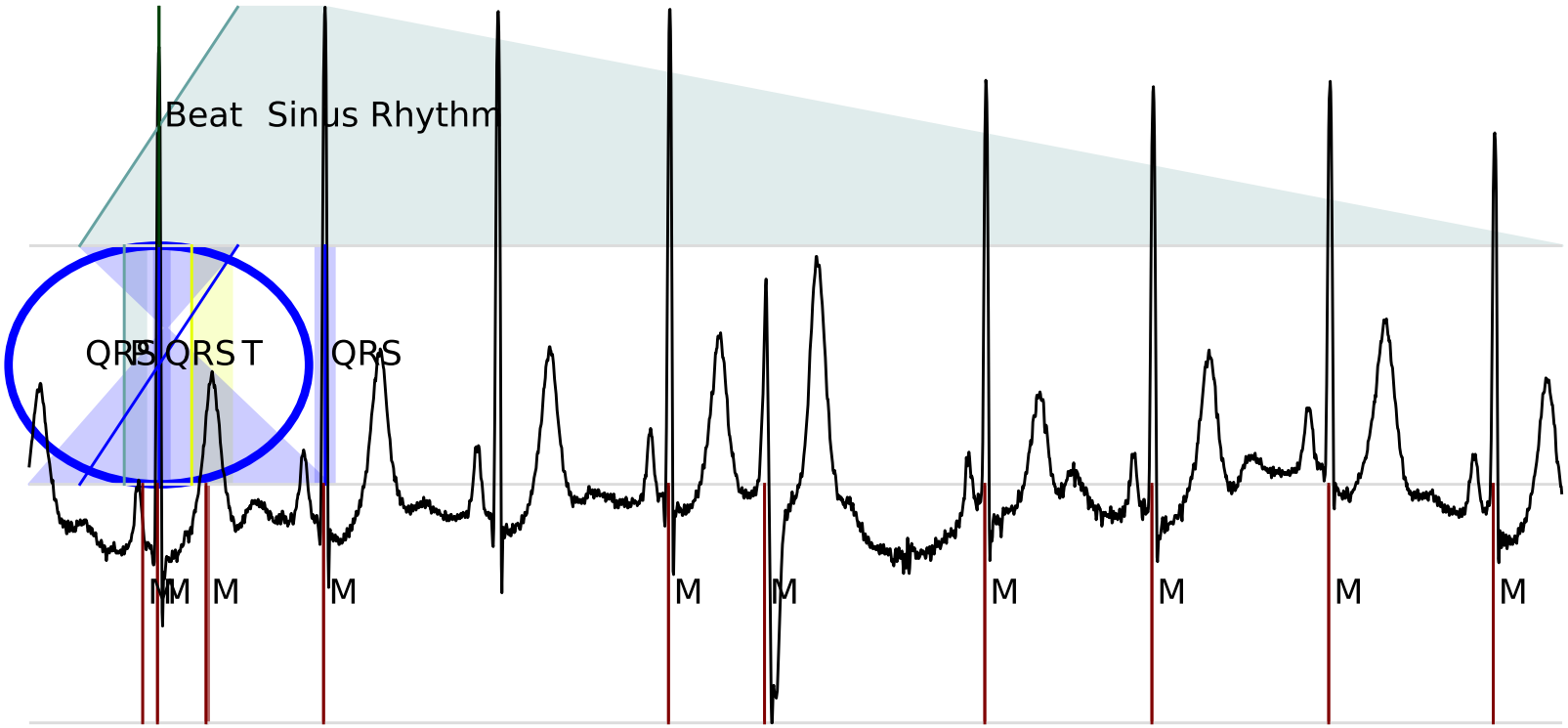
Interpretation example



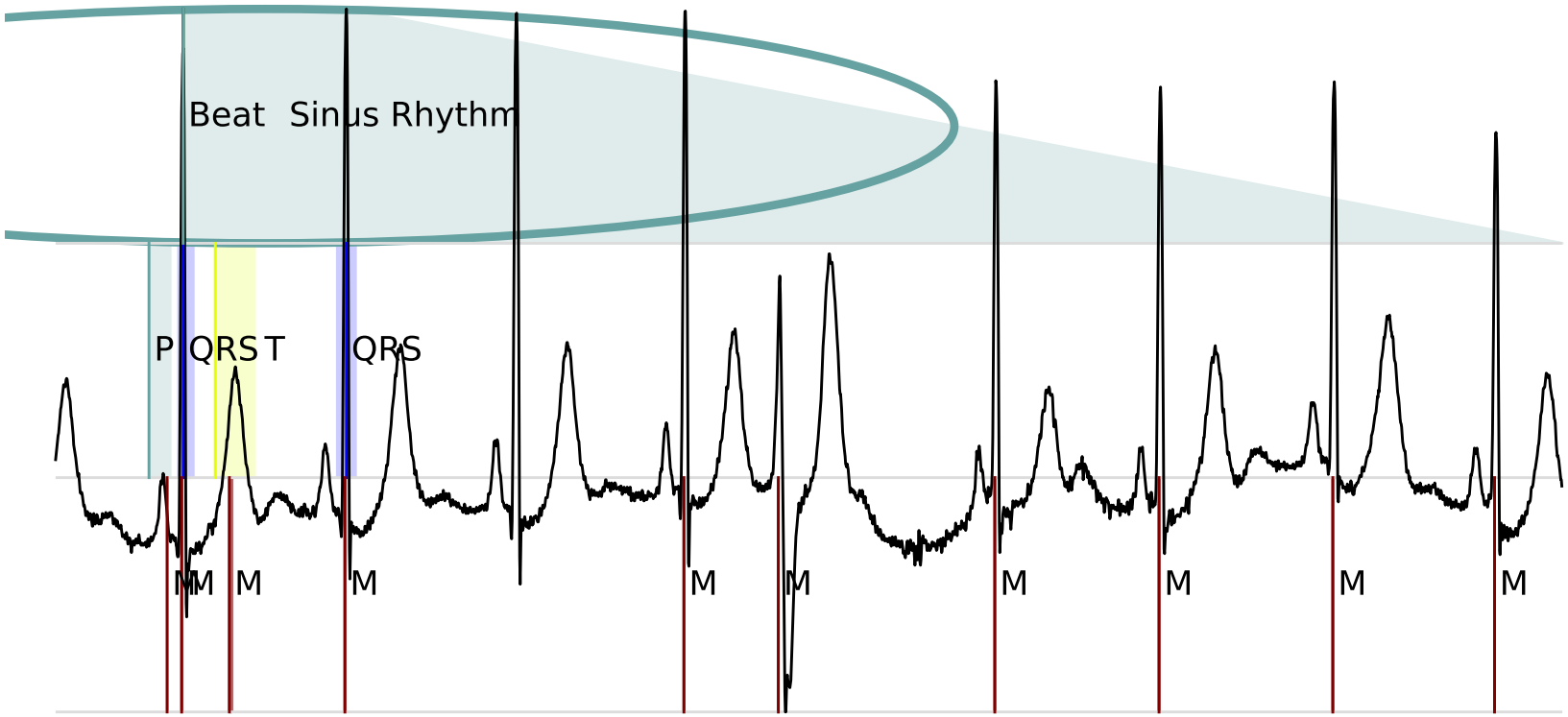
Interpretation example



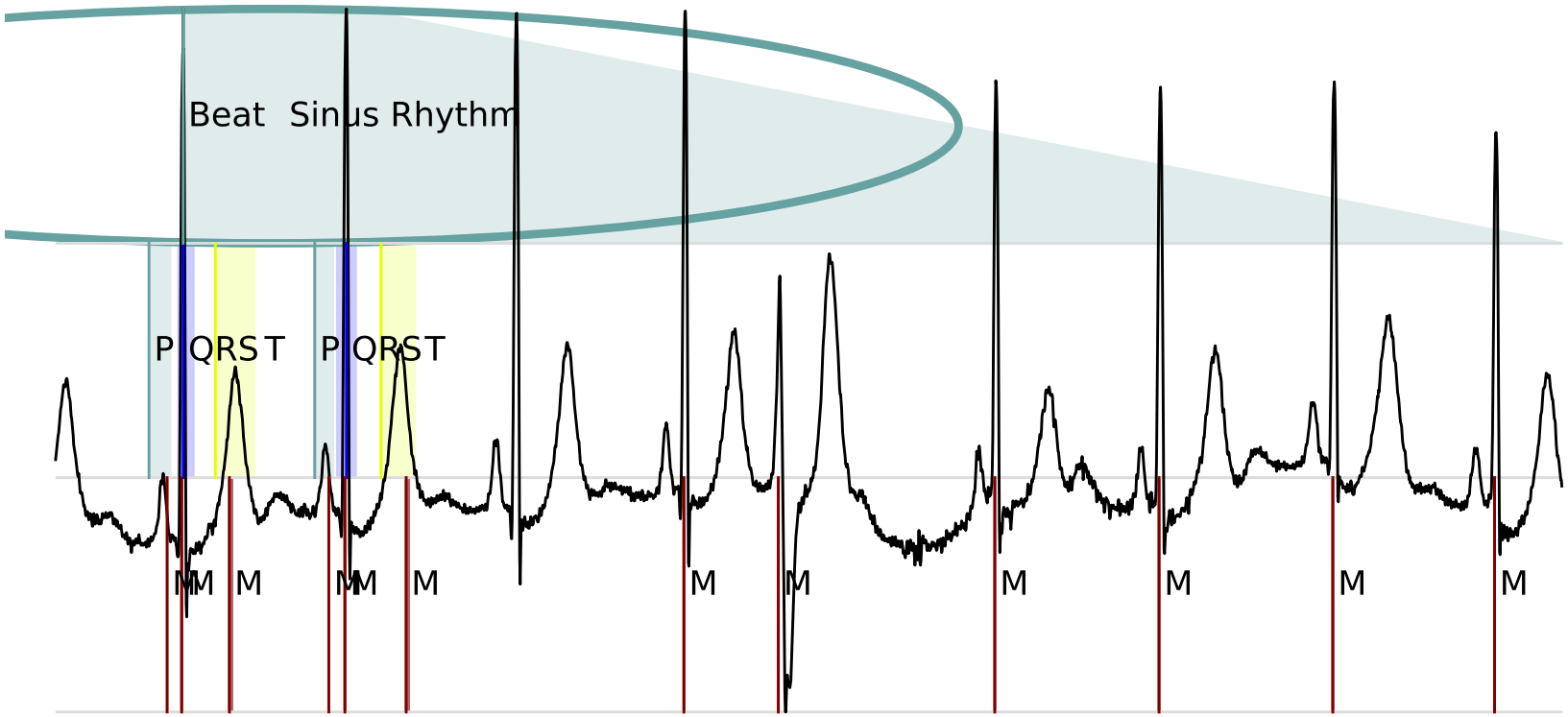
Interpretation example



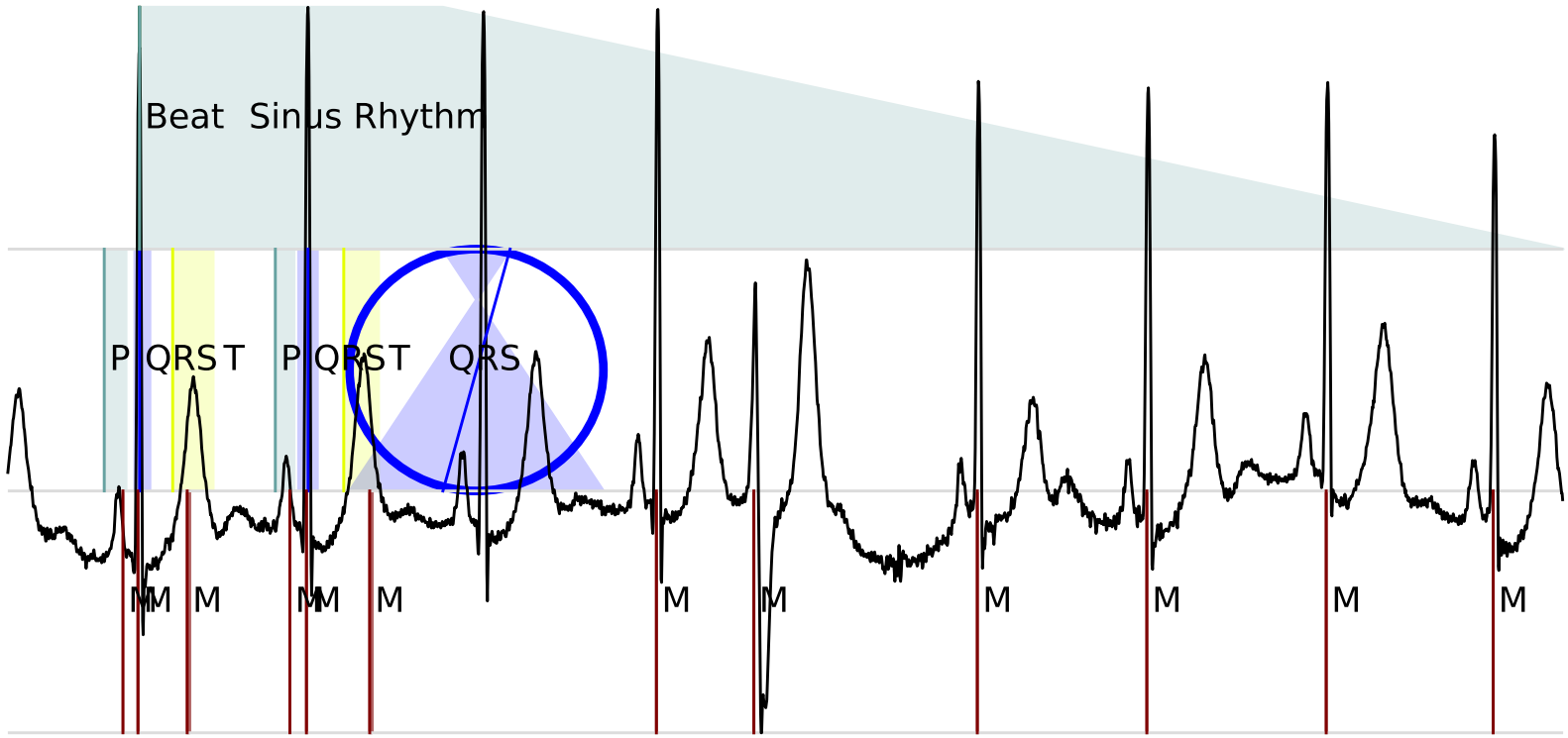
Interpretation example



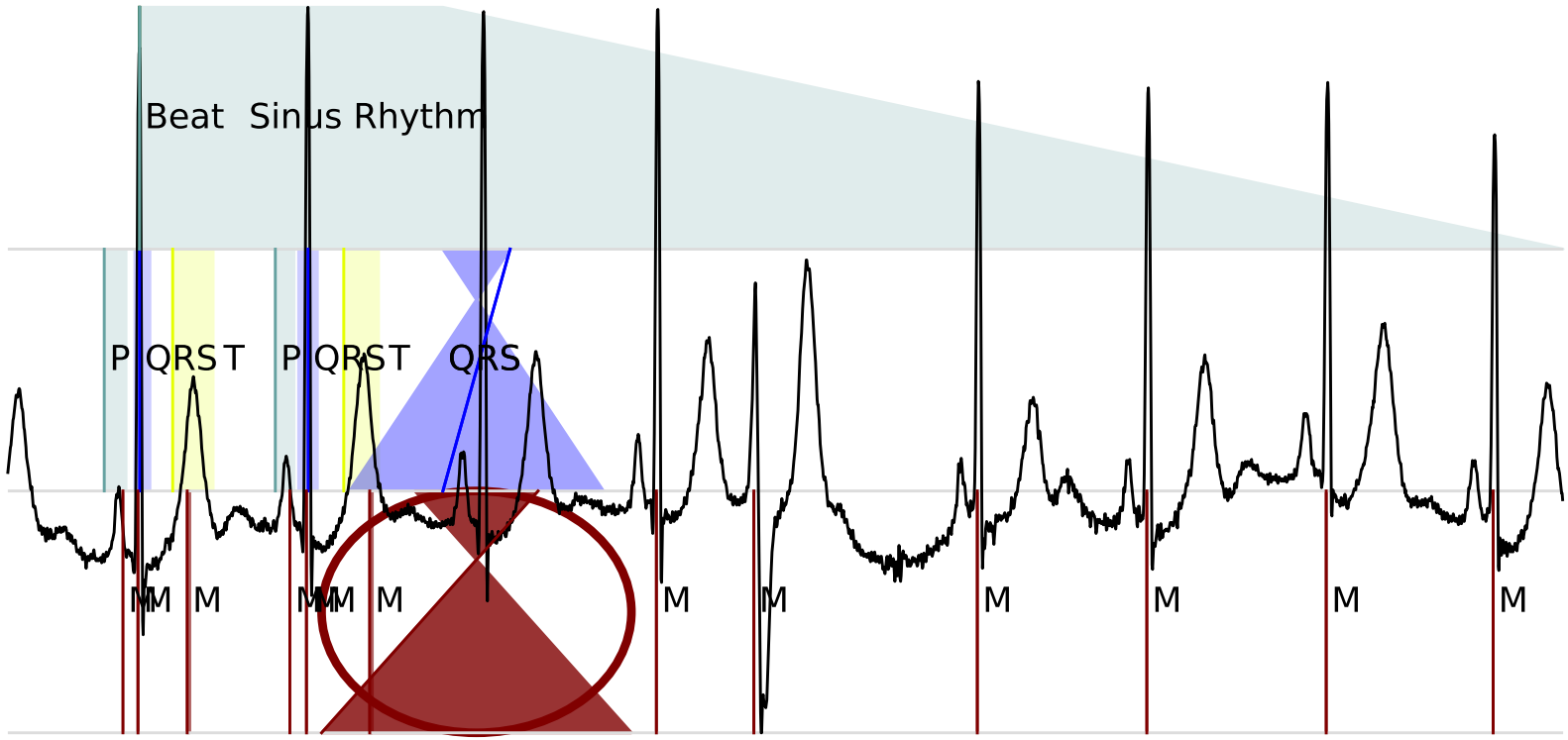
Interpretation example



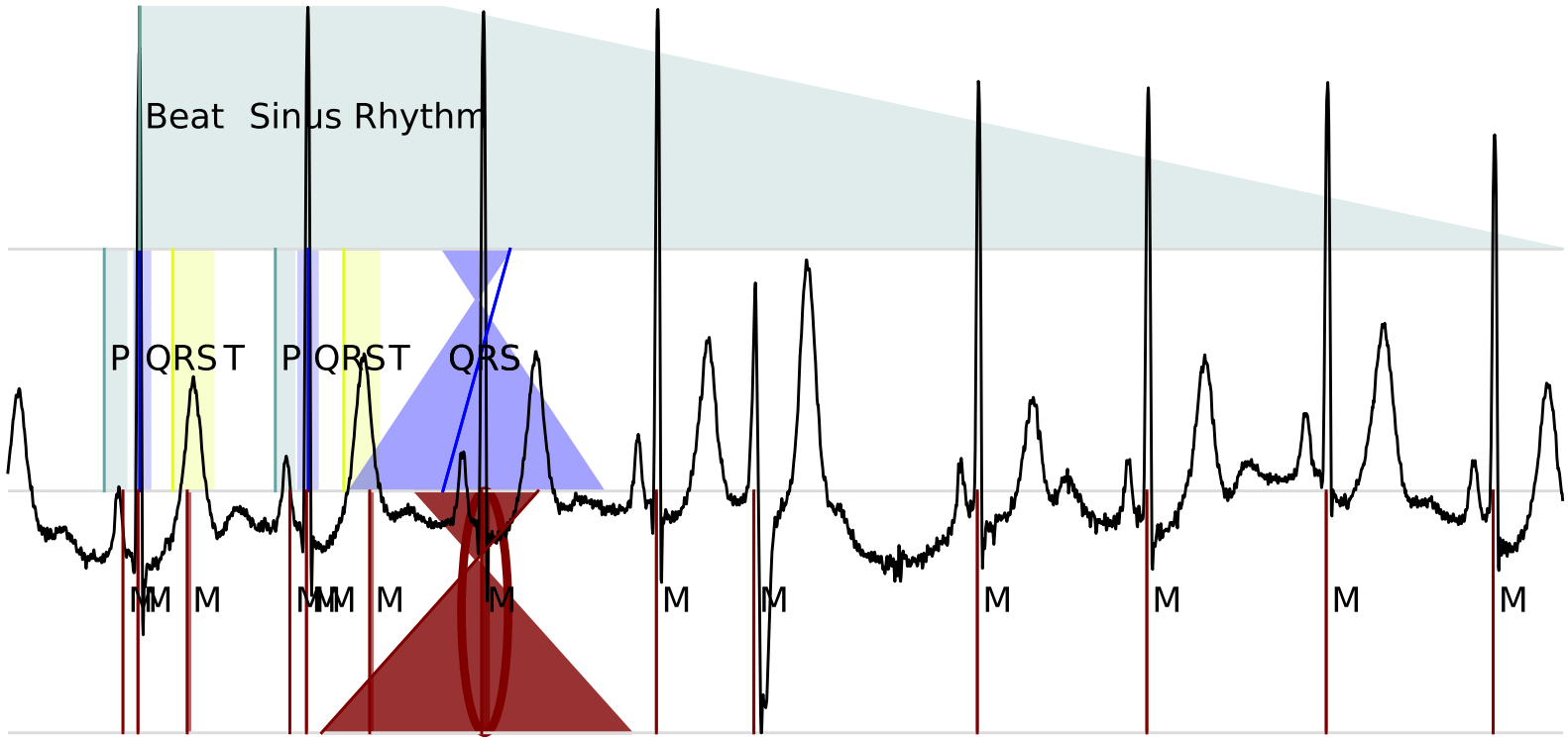
Interpretation example



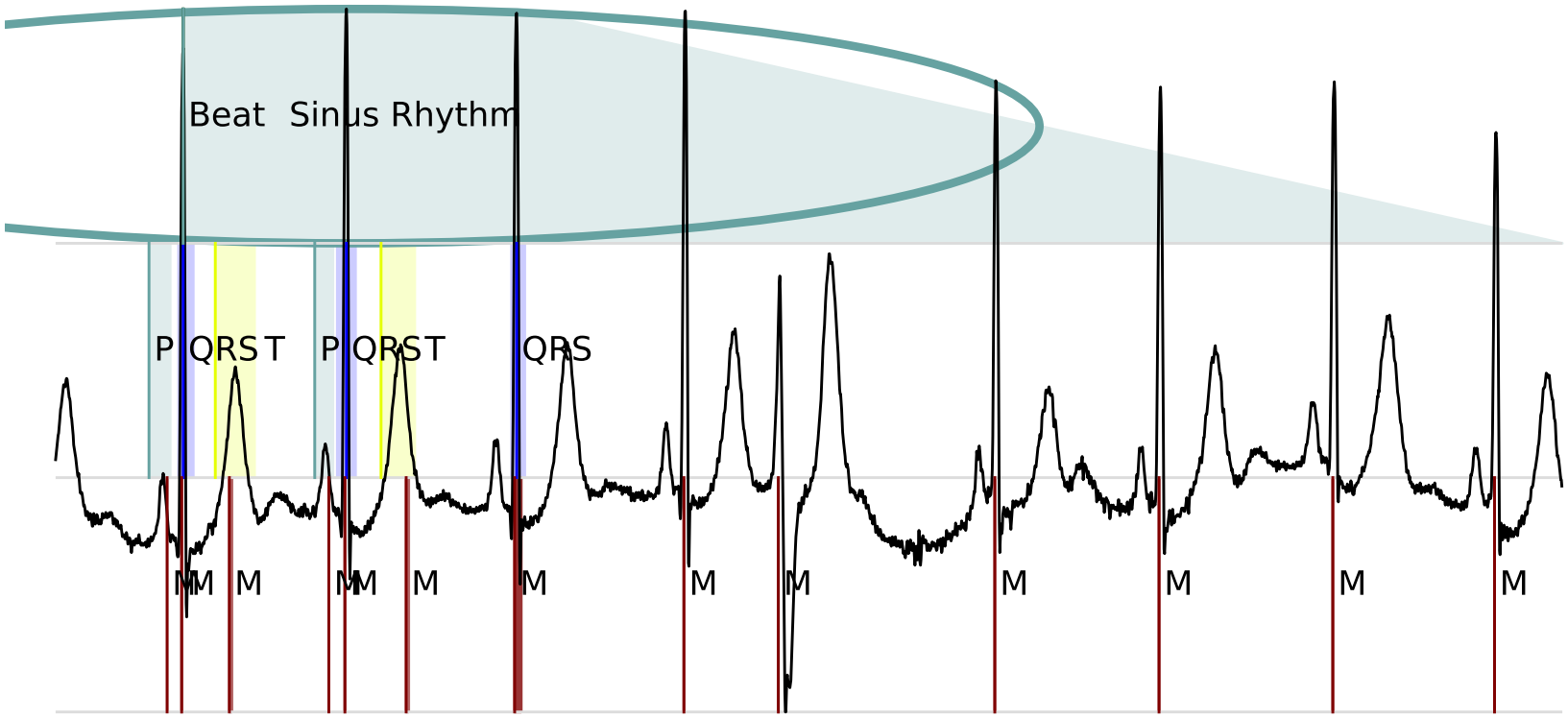
Interpretation example



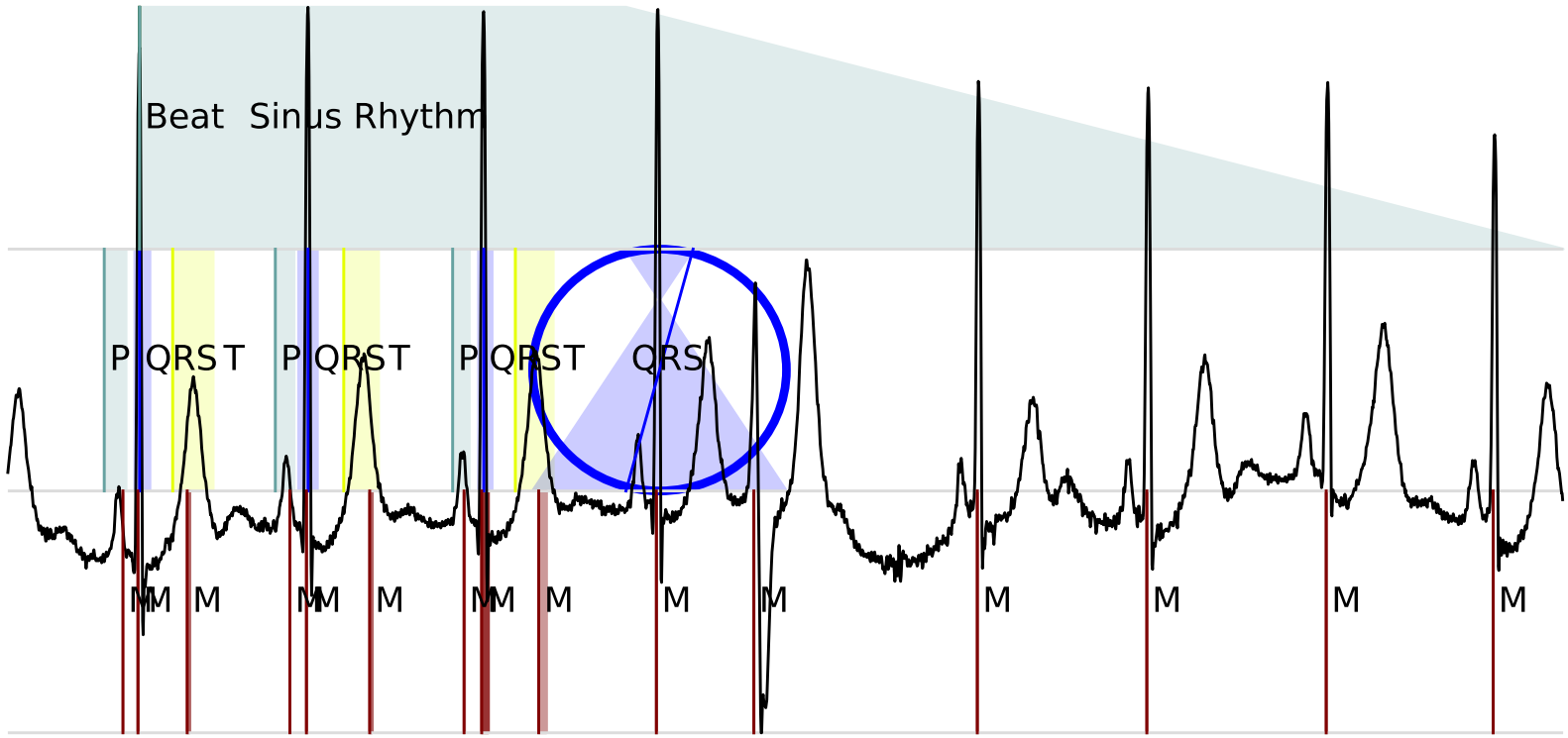
Interpretation example



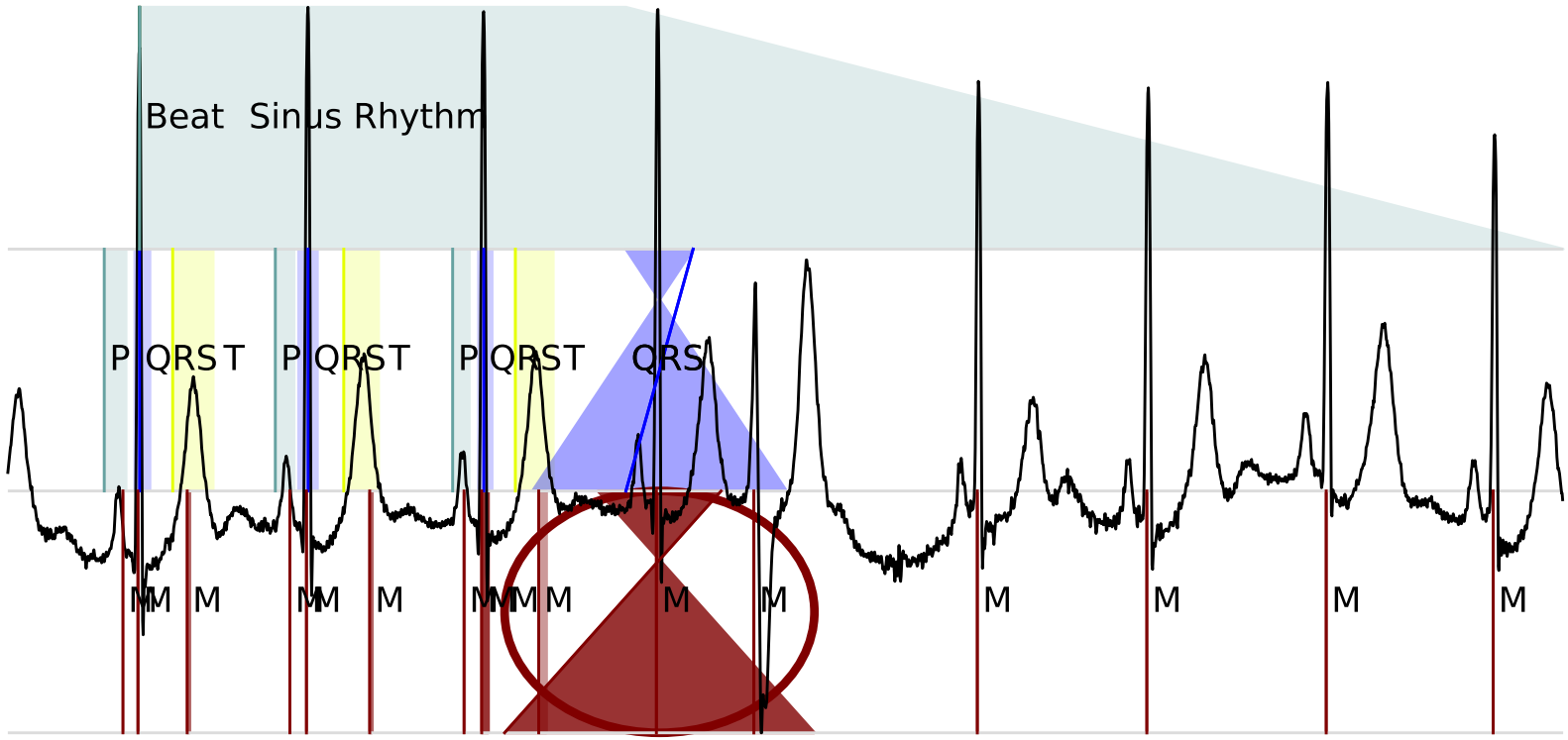
Interpretation example



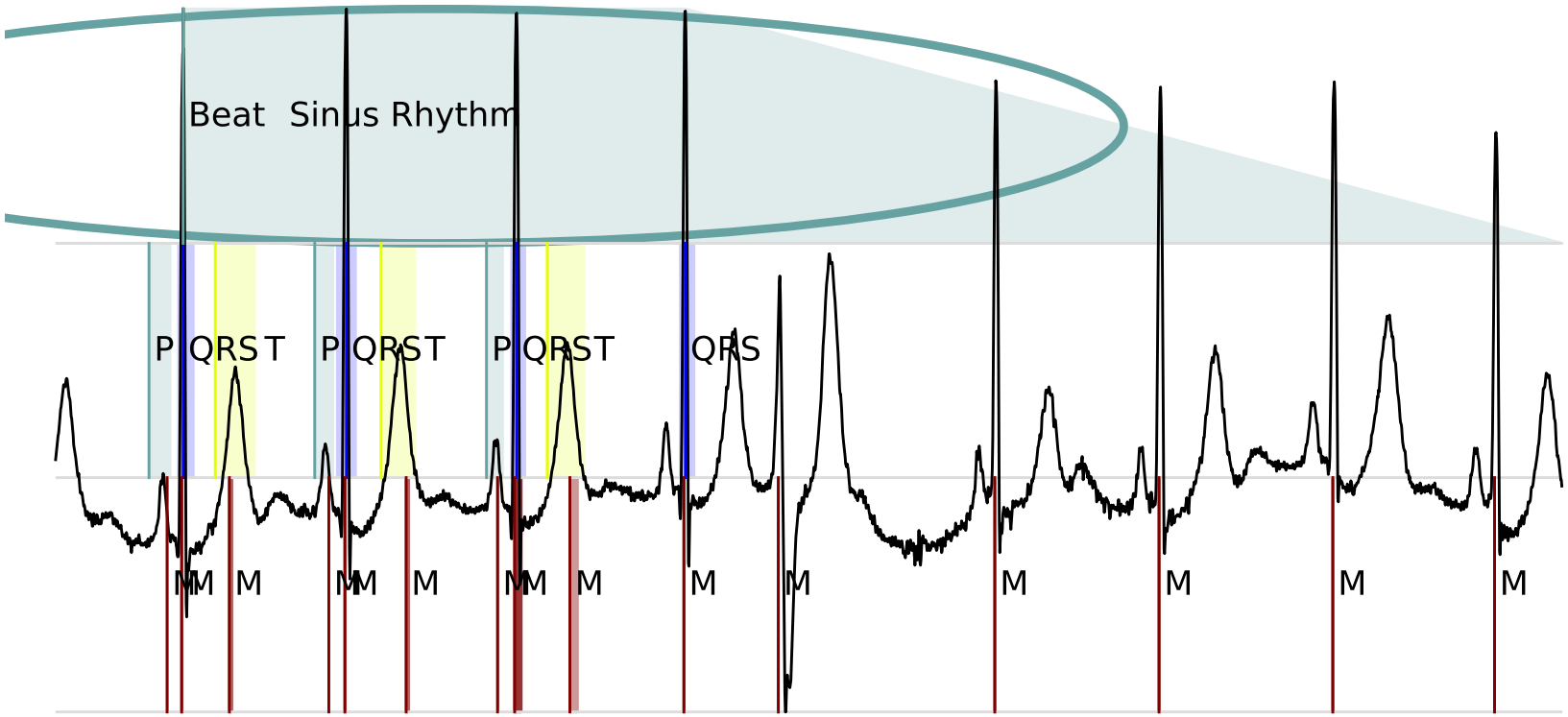
Interpretation example



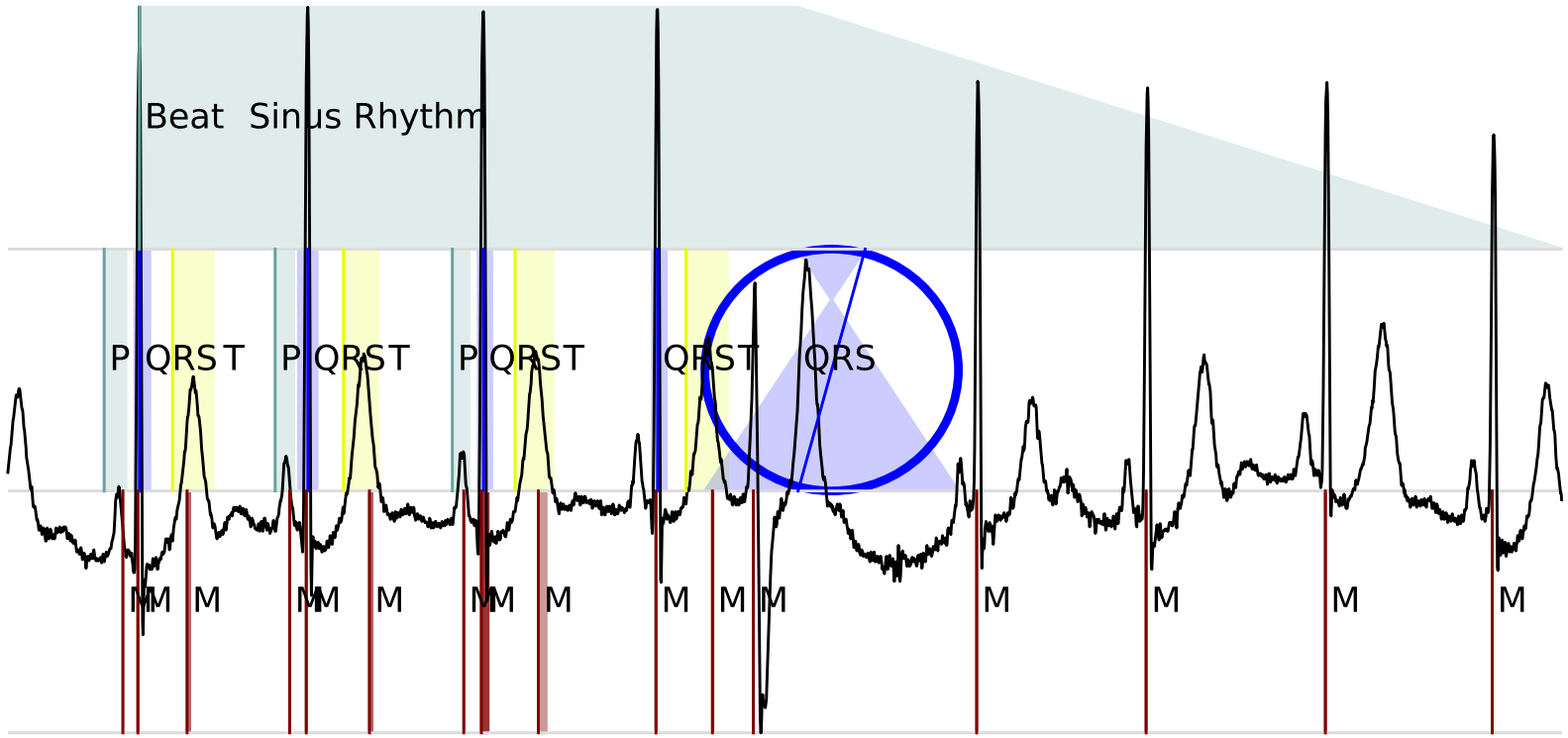
Interpretation example



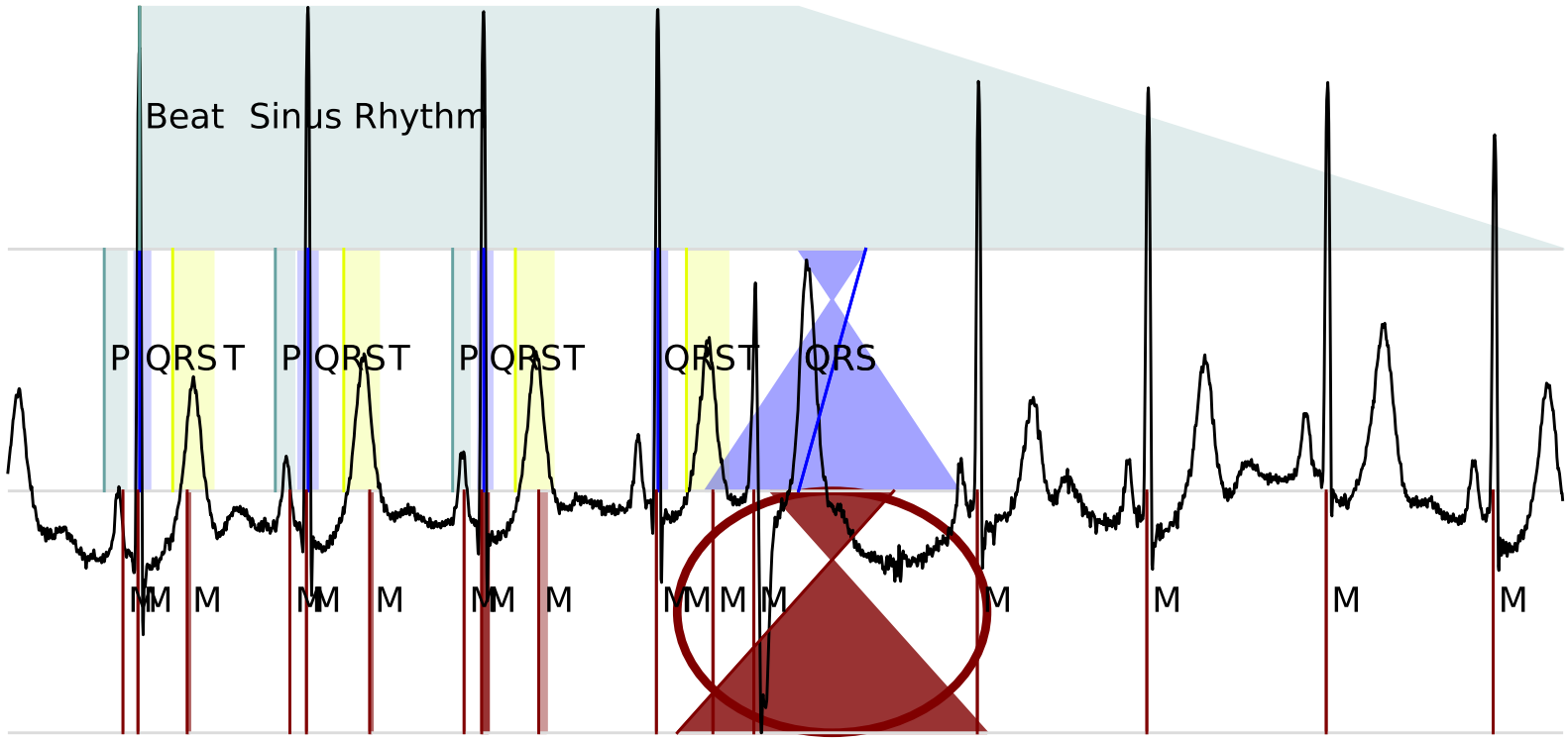
Interpretation example



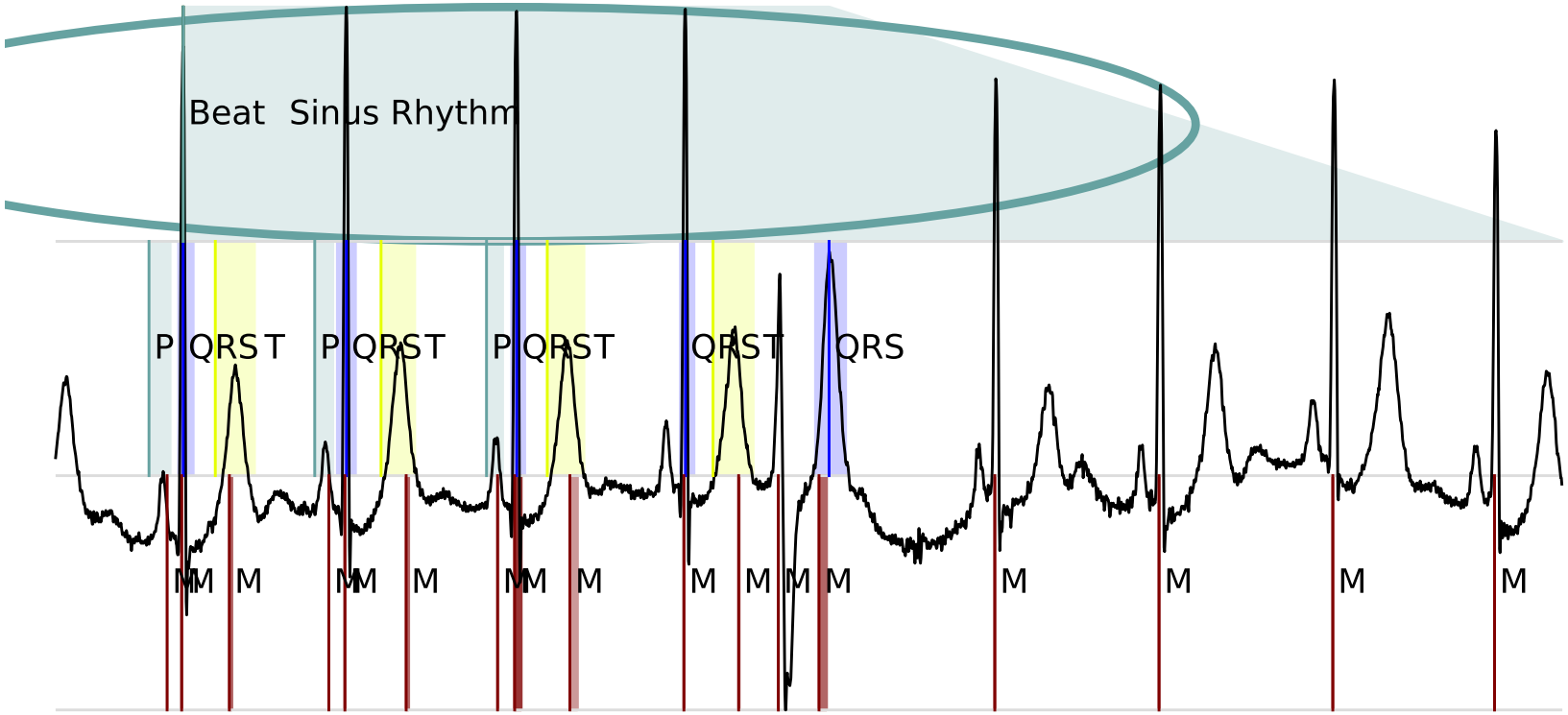
Interpretation example



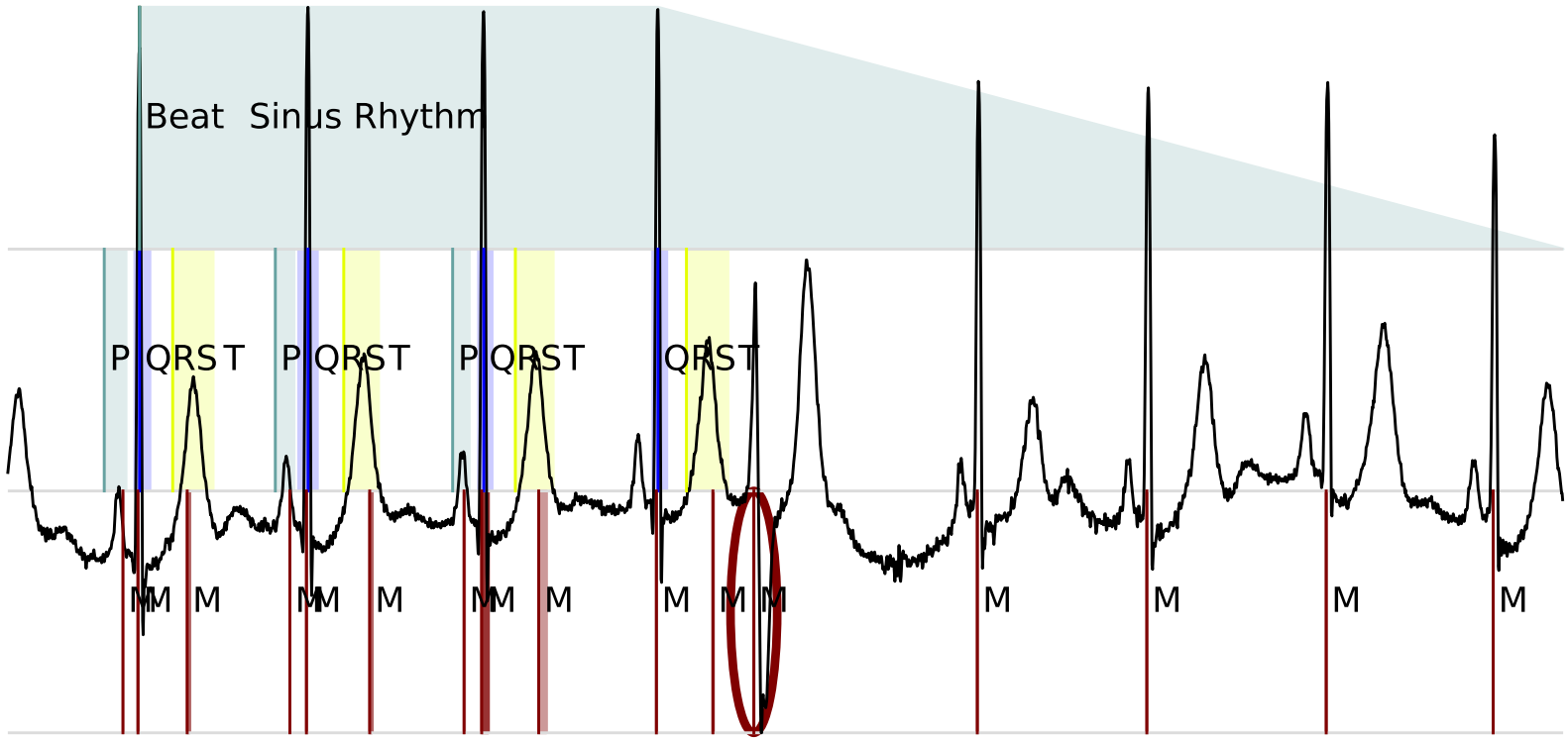
Interpretation example



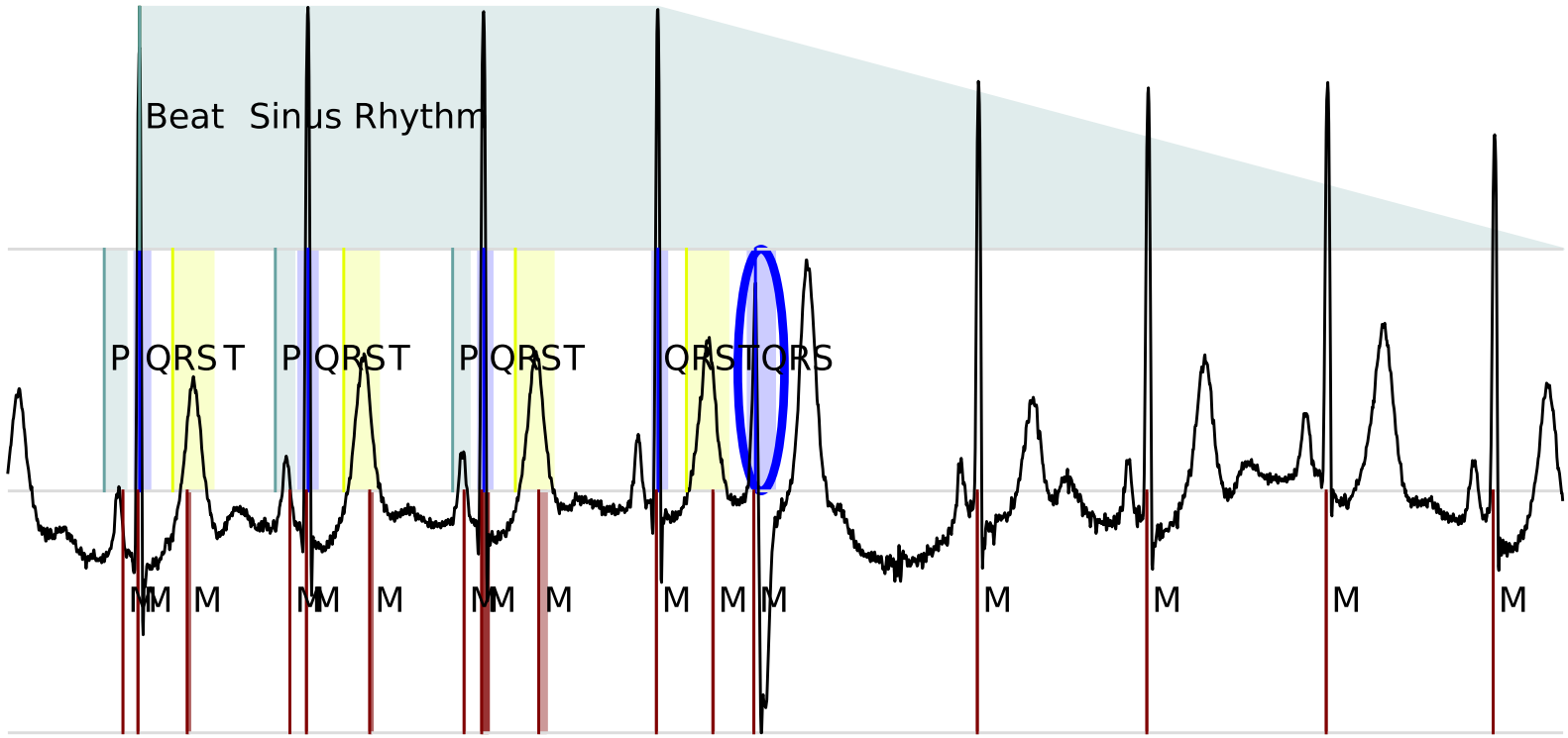
Interpretation example



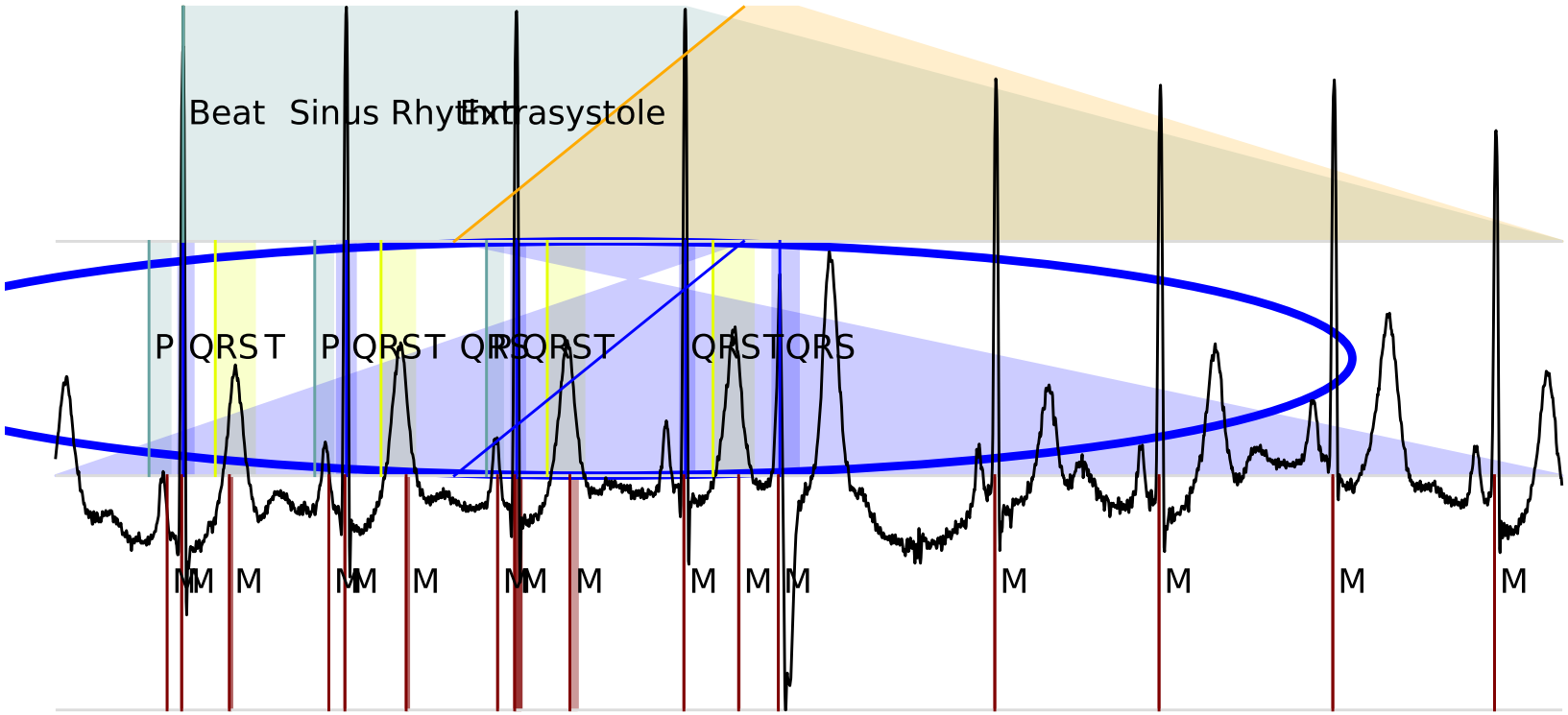
Interpretation example



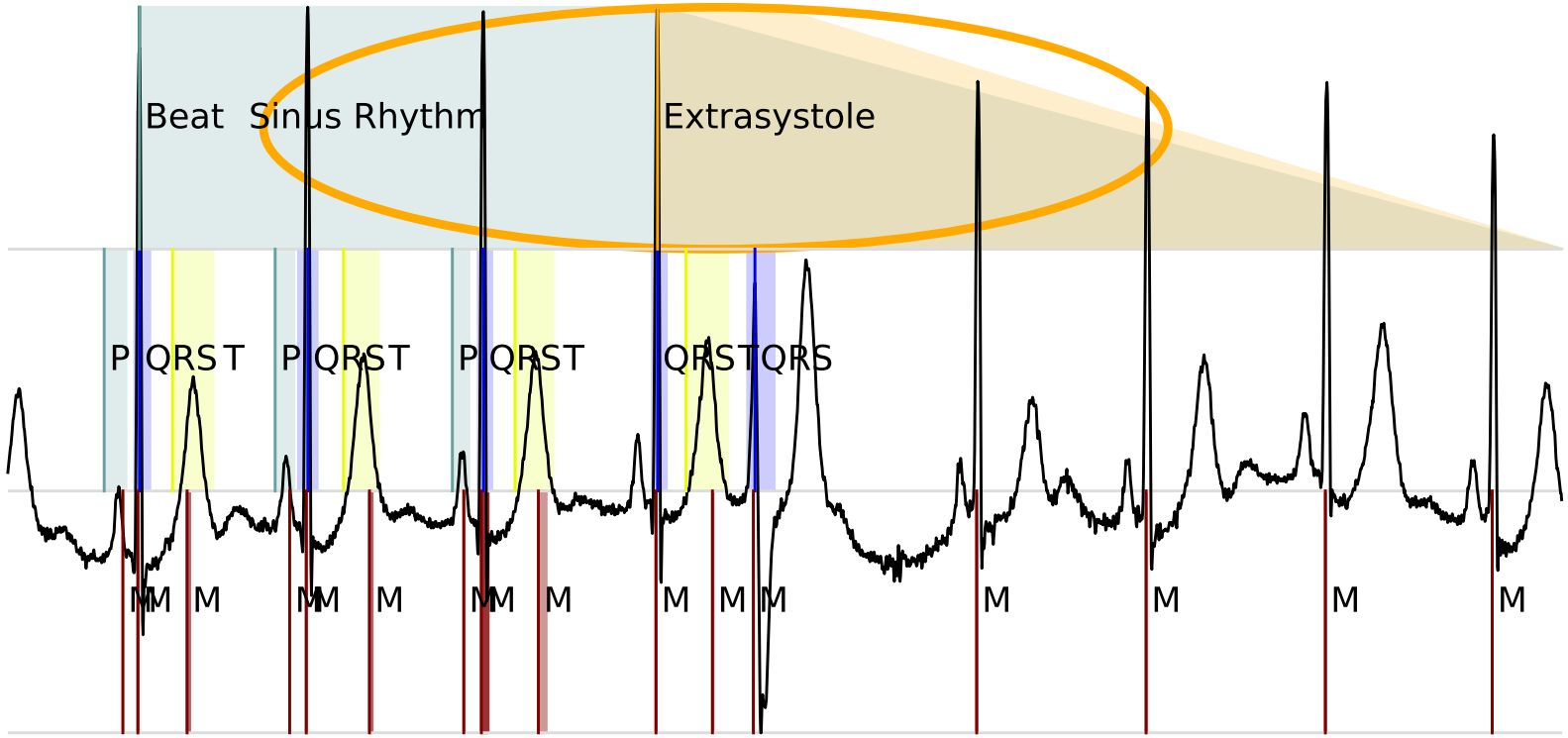
Interpretation example



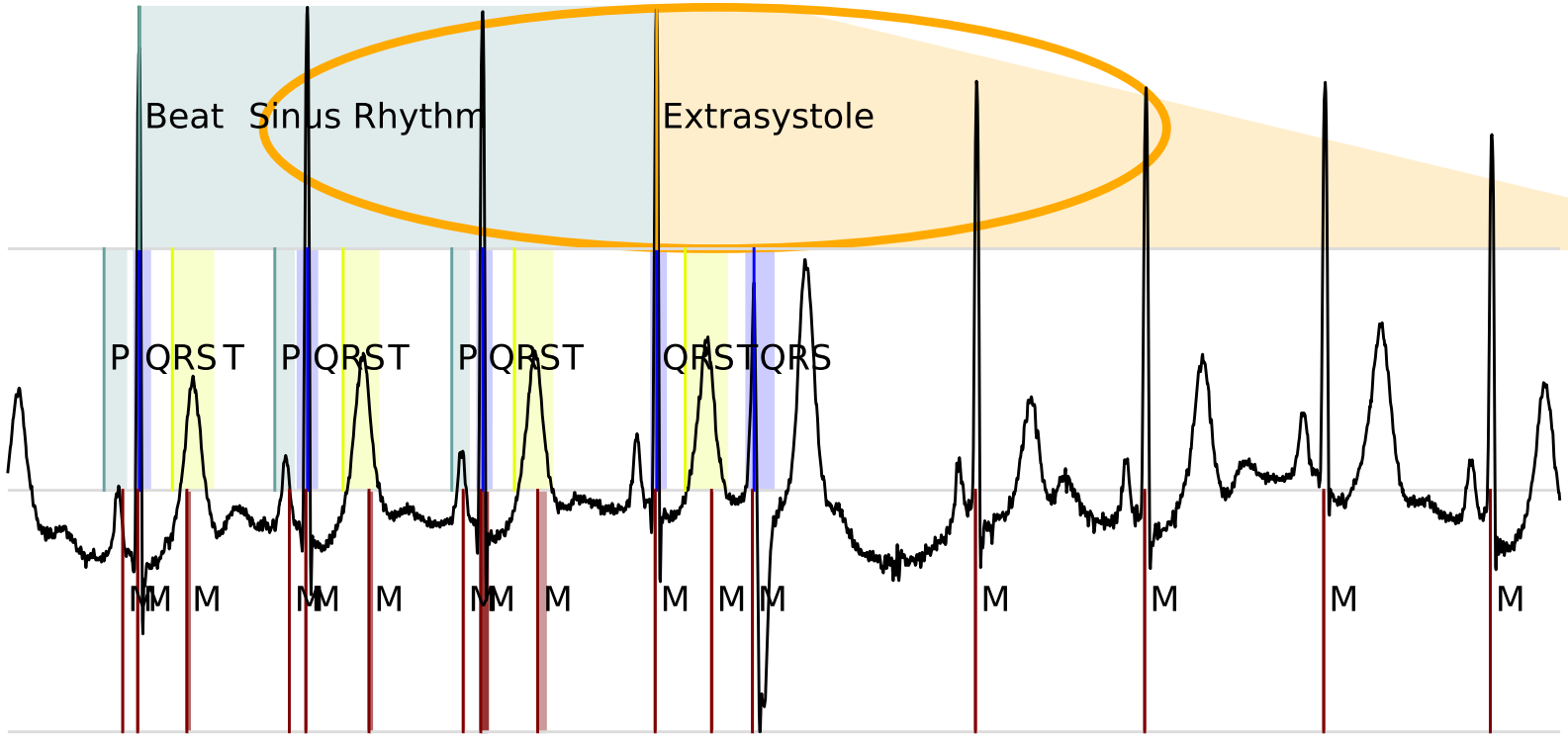
Interpretation example



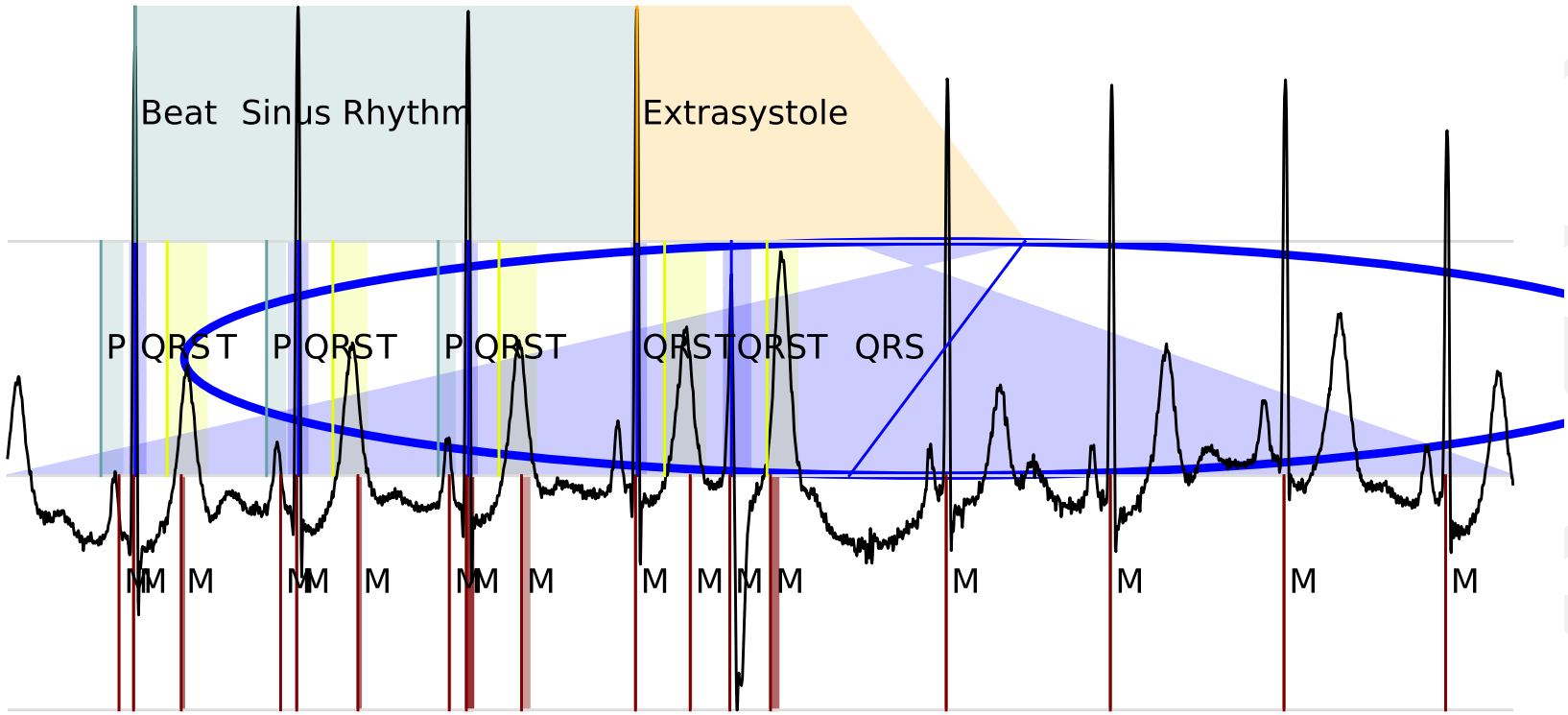
Interpretation example



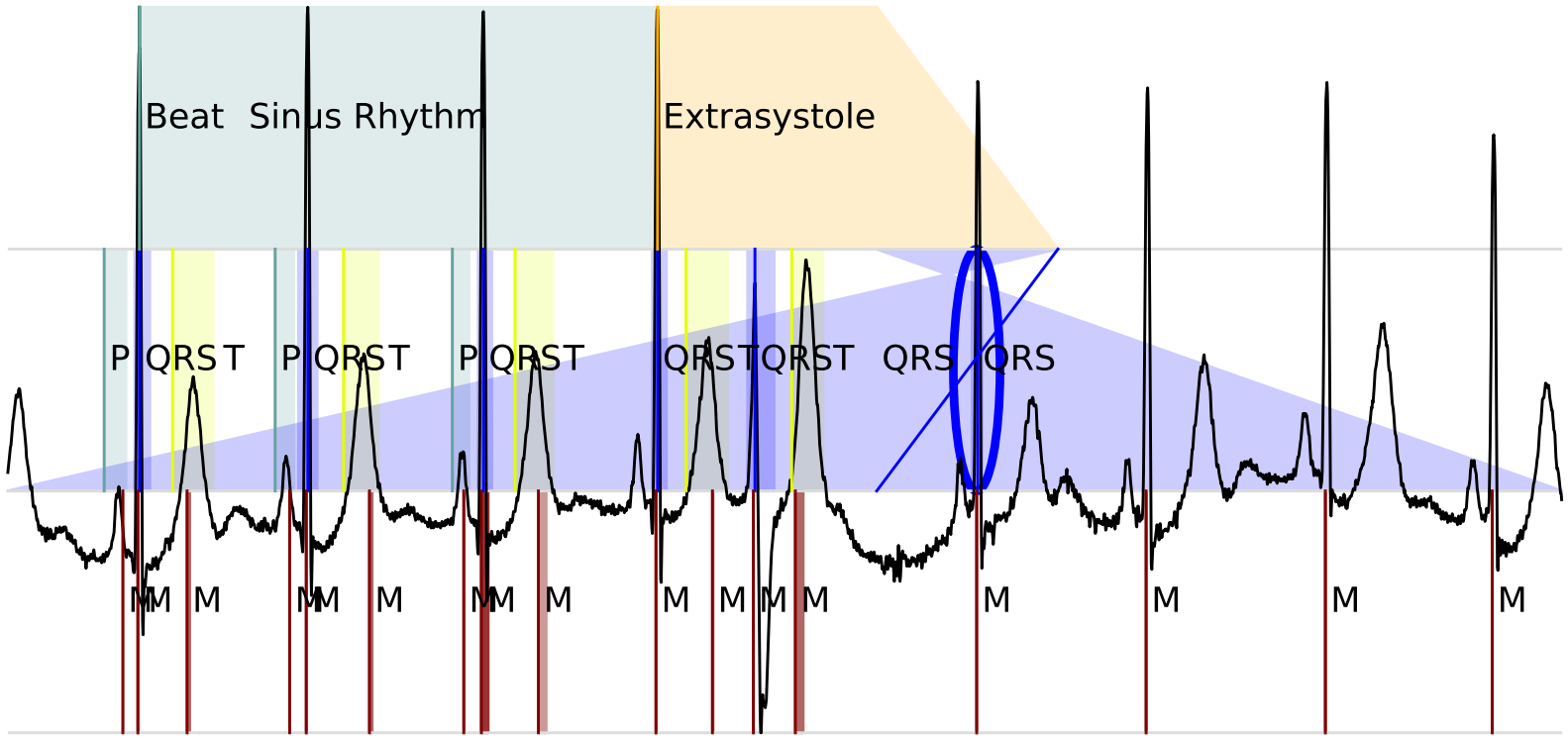
Interpretation example



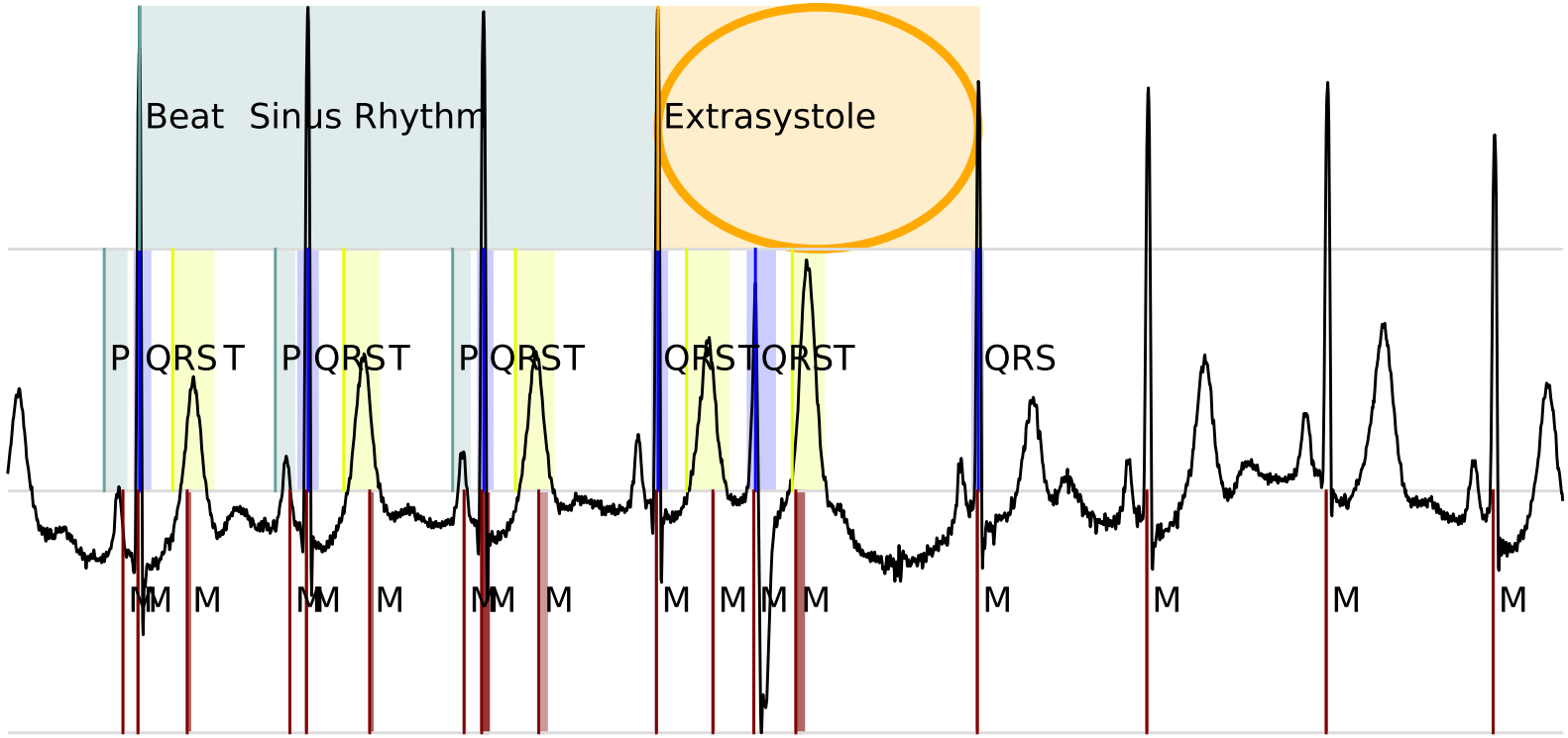
Interpretation example



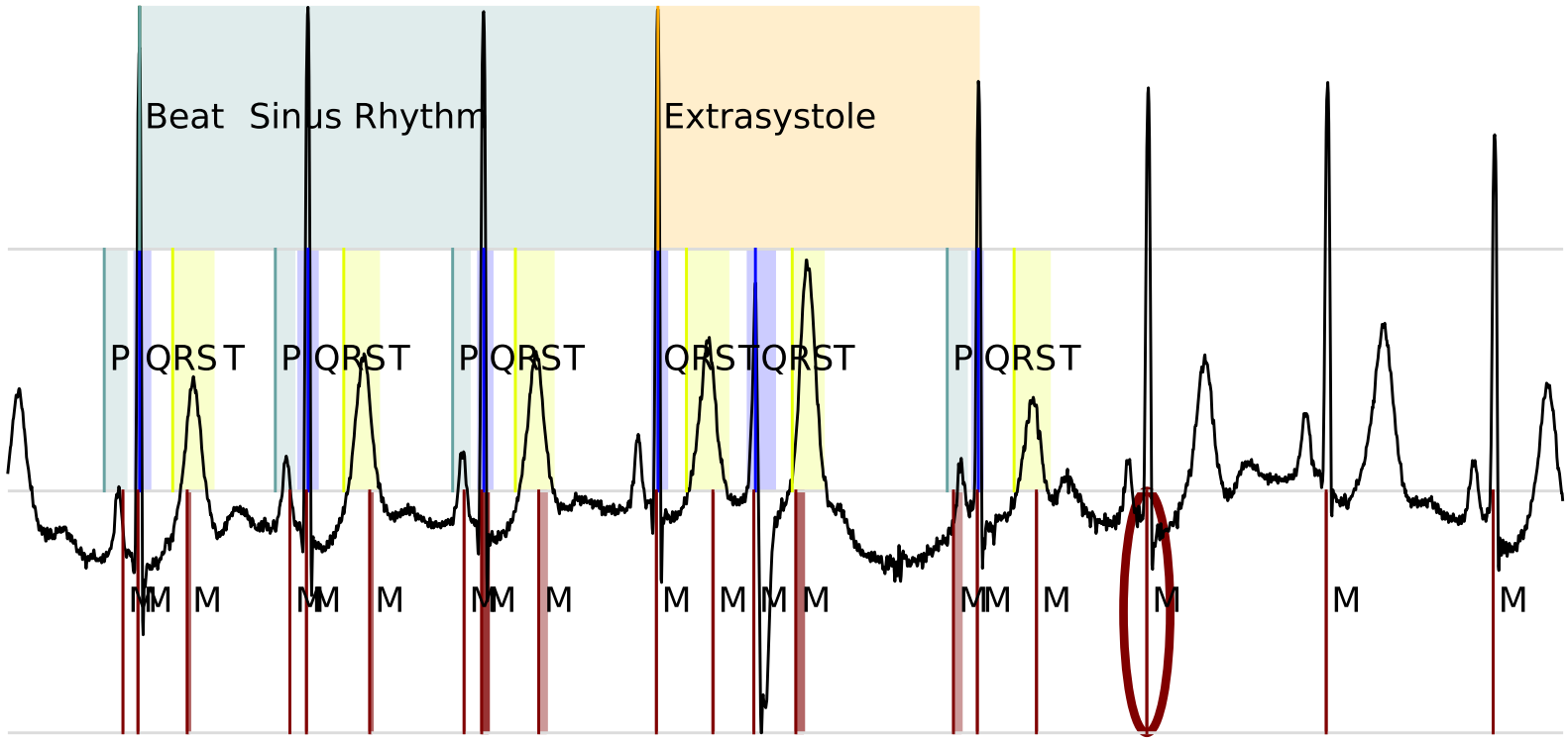
Interpretation example



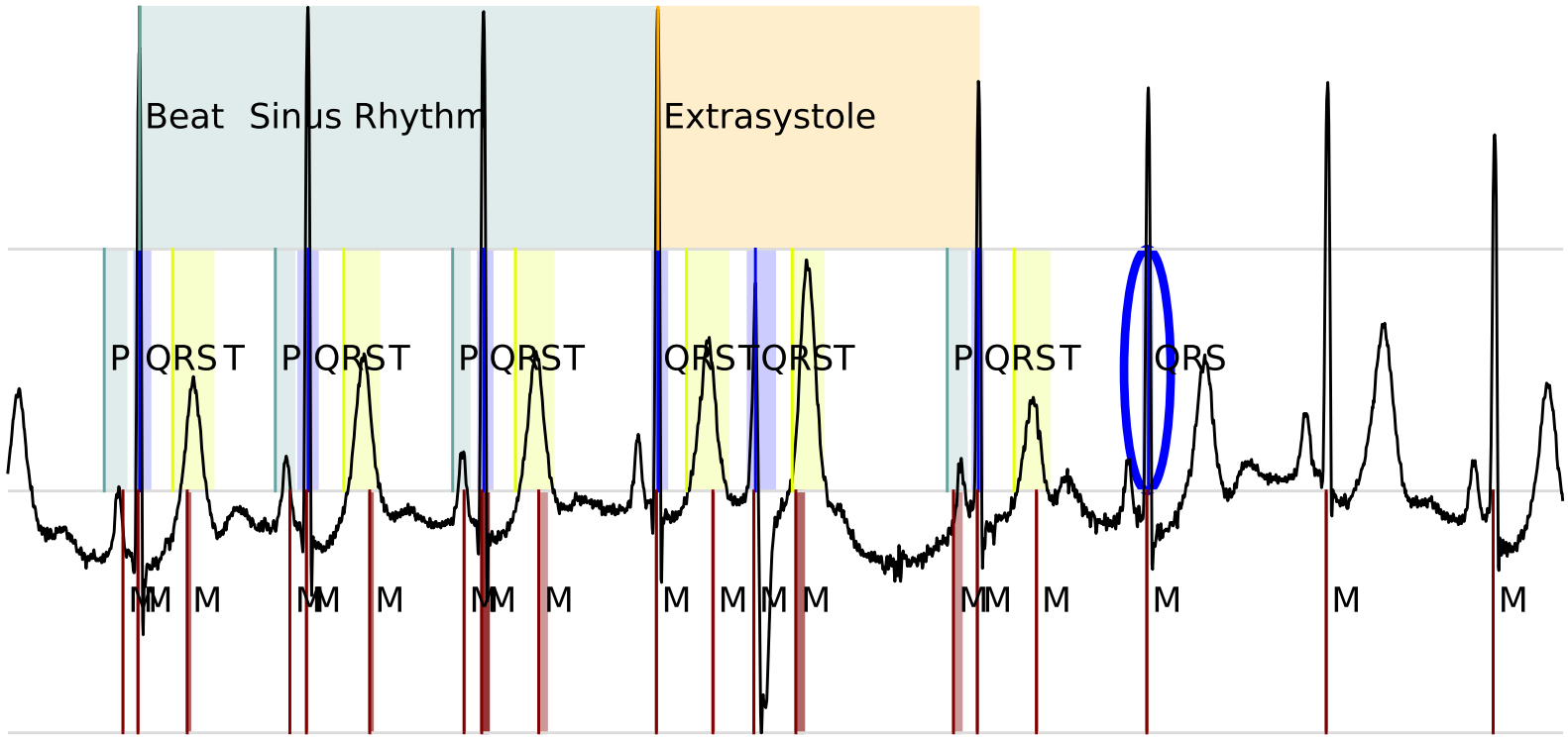
Interpretation example



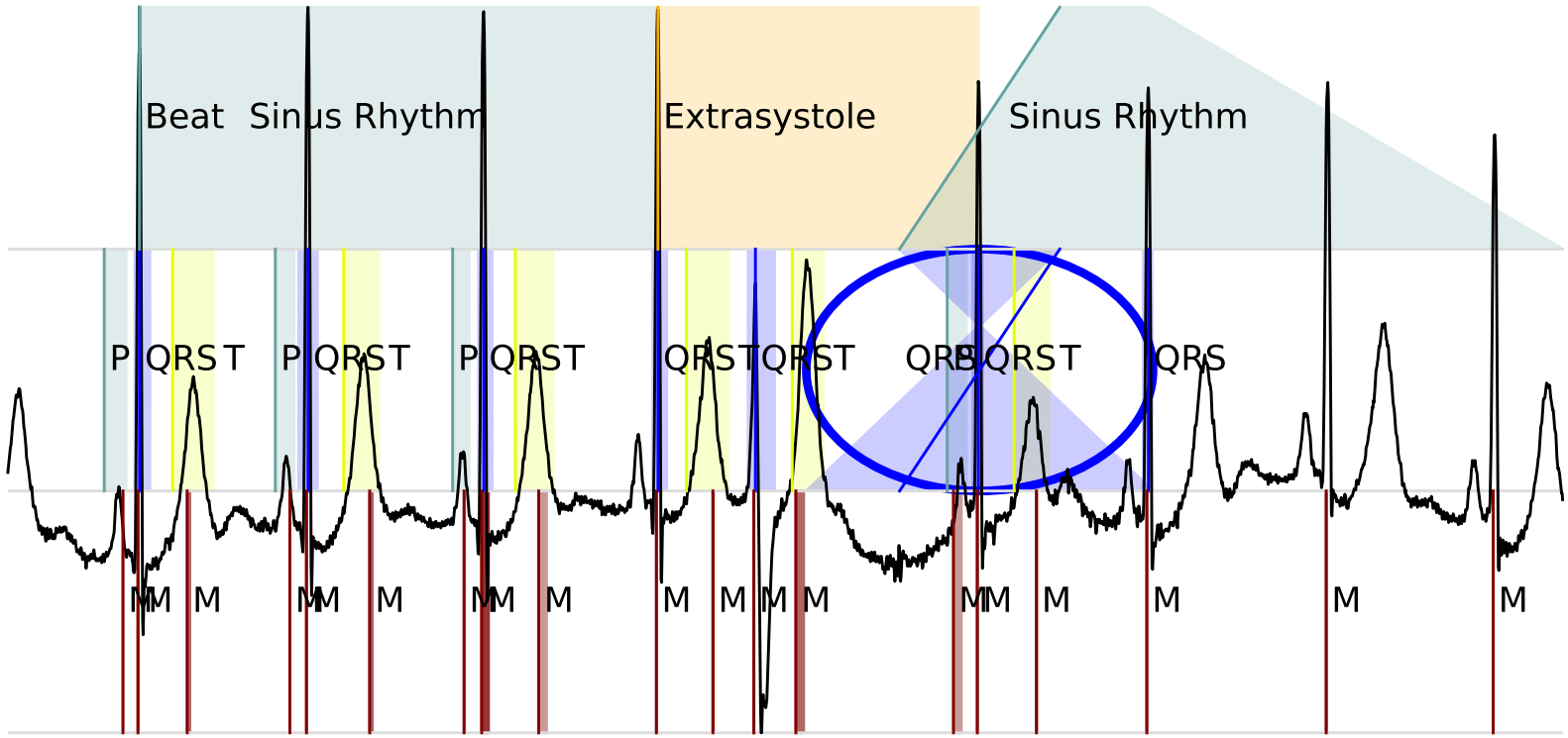
Interpretation example



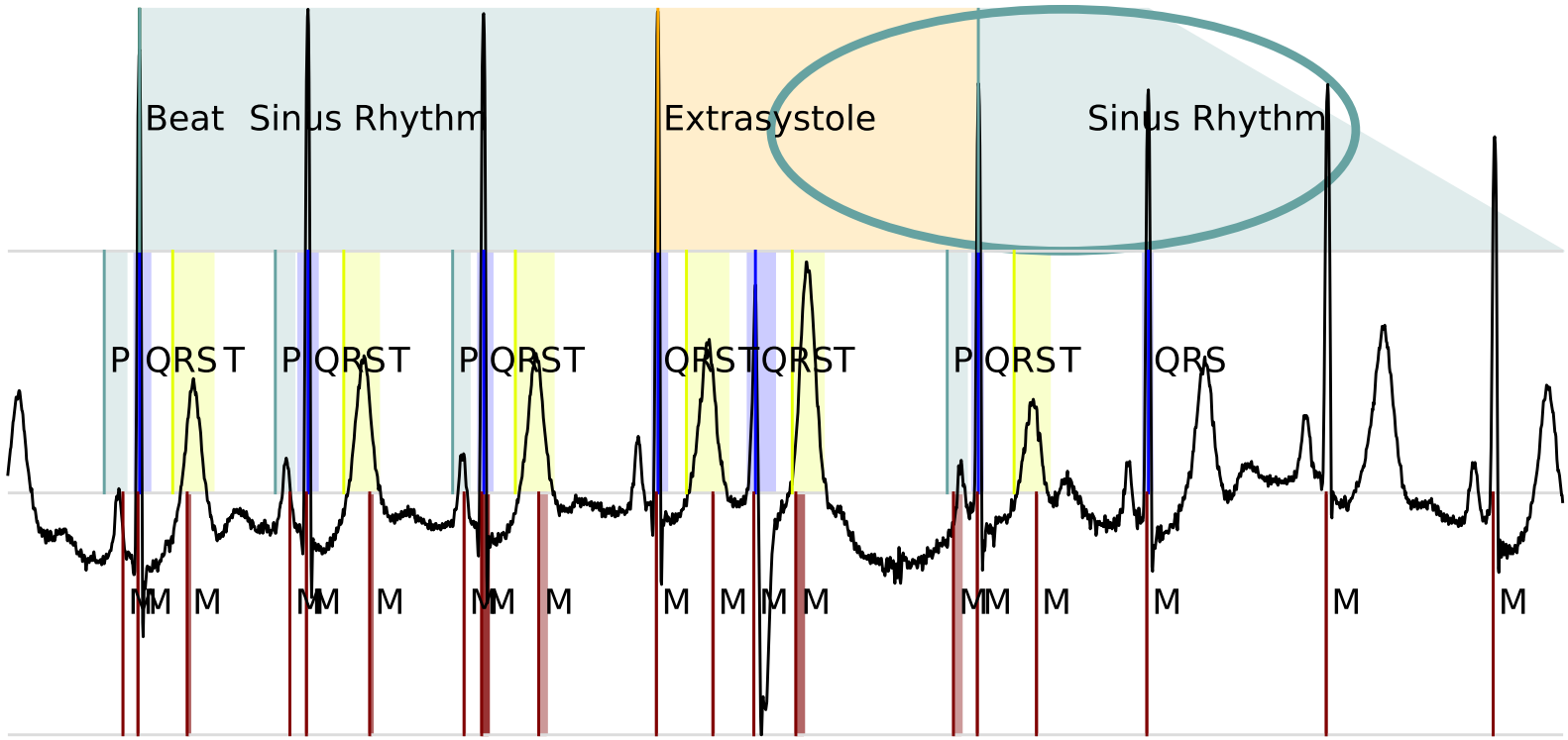
Interpretation example



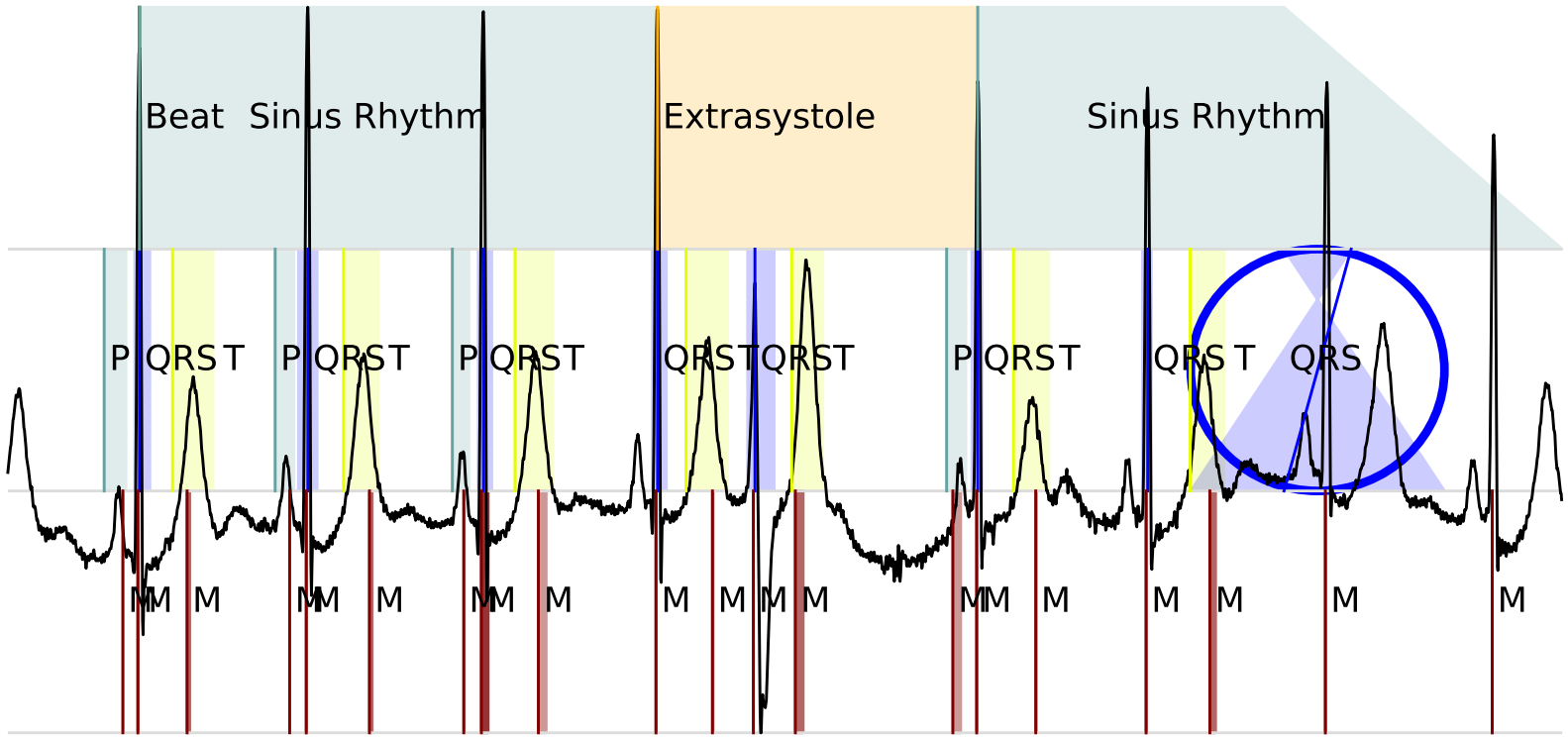
Interpretation example



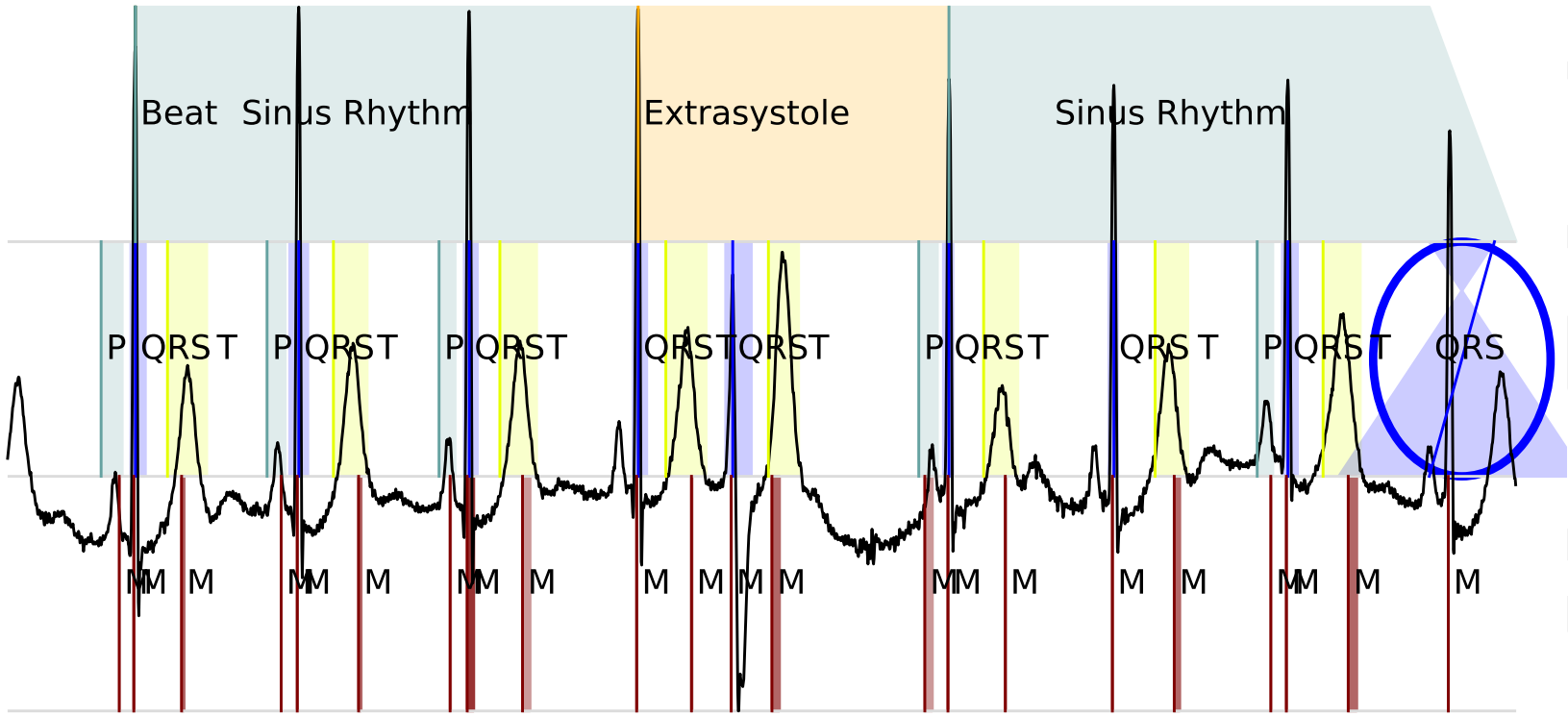
Interpretation example



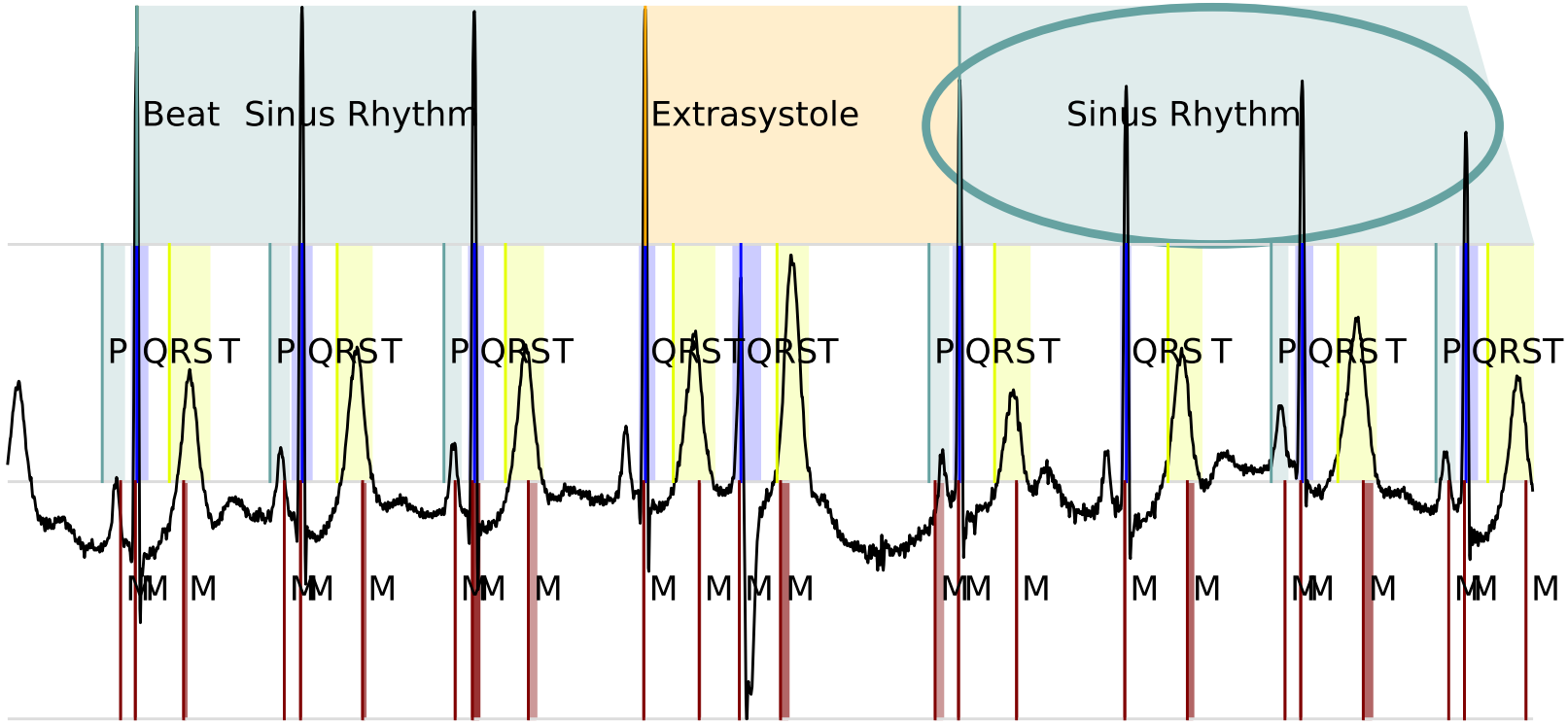
Interpretation example



Interpretation example

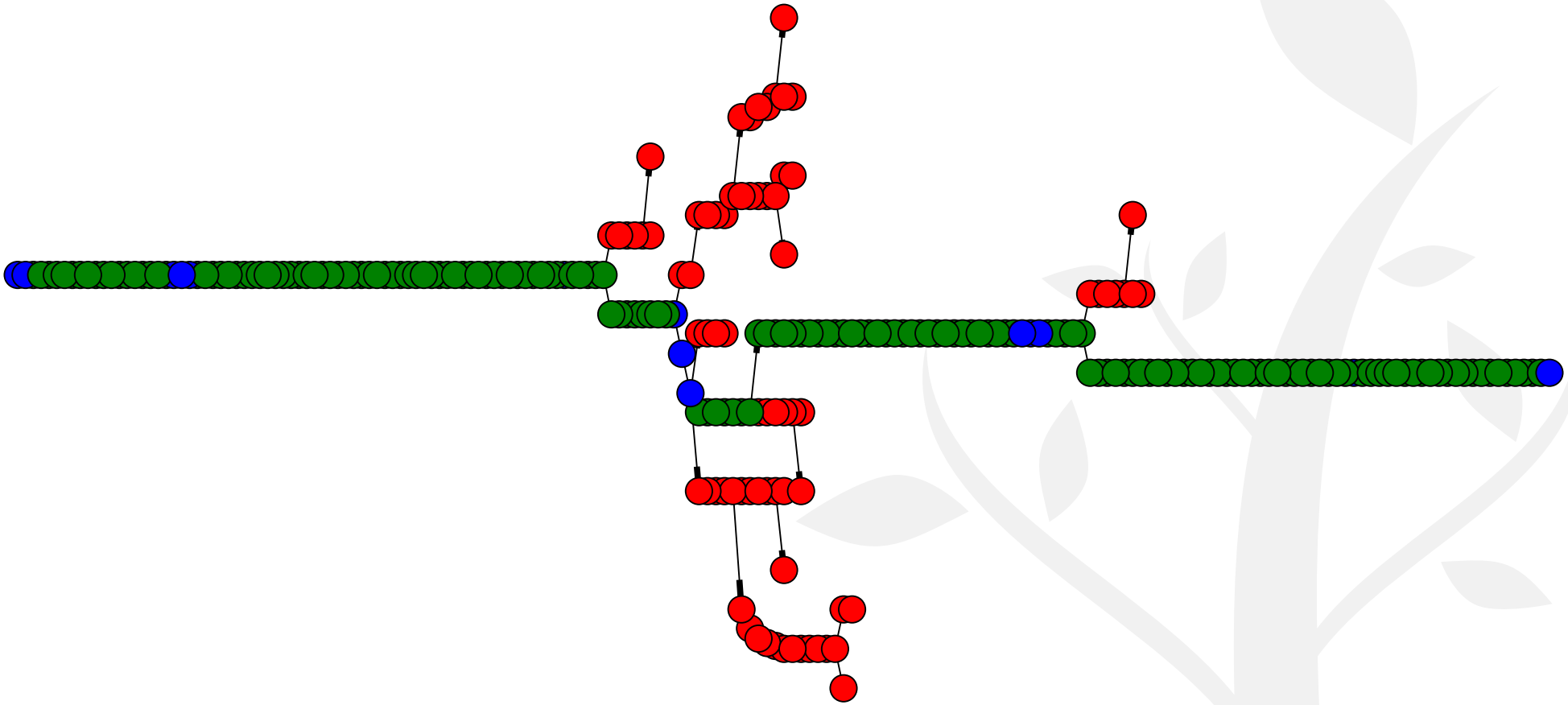


Interpretation example



Interpretation example

Interpretation search graph:



Validation results

Experiment

Improve an state of the art QRS detector by including some basic rhythm abstraction patterns:

- *Normal Rhythm*: Regular heart-rate in 60-100 bpm.
- *Tachycardia*: Regular heart-rate in 100-250 bpm.
- *Bradycardia*: Regular heart-rate in 30-60 bpm.
- *Extrasystole*: Advanced beat with a posterior compensatory pause.

Results with records from two databases, accumulating over 1 700 000 beats:

	WQRS			WQRS + Abduction			
Database	**Se**	**P+**	**F1**	**Se**	**P+**	**F1**	**P-value**
NSR	99.90	99.08	99.49	99.83	99.40	**99.61**	**0.008**
MIT-BIH	99.95	98.58	99.26	99.62	99.64	**99.63**	**0.033**

Future work

- Define a **comprehensive knowledge base** for interpreting a wider range of phenomena and that can address more complex problems, such as beat classification by origin and arrhythmia classification.
- Evaluate the approach in **multimodal records** (more than one type of signal).
- Include an **hypotheses evaluation scheme** (probability calculus, for example).

Future work

- Define a **comprehensive knowledge base** for interpreting a wider range of phenomena and that can address more complex problems, such as beat classification by origin and arrhythmia classification.
- Evaluate the approach in **multimodal records** (more than one type of signal).
- Include an **hypotheses evaluation scheme** (probability calculus, for example).

Future work

- Define a **comprehensive knowledge base** for interpreting a wider range of phenomena and that can address more complex problems, such as beat classification by origin and arrhythmia classification.
- Evaluate the approach in **multimodal records** (more than one type of signal).
- Include an **hypotheses evaluation scheme** (probability calculus, for example).

Thank you very much for your attention!

Questions?