Developing Software for Fully Automatic Creation of Accessible SVG Diagrams

John Gardner

ViewPlus Technologies
Corvallis, Oregon
viewplus.com

Volker Sorge

Progressive Accessibility Solutions
Birmingham, UK
progressiveaccess.com

CSUN 2017, San Diego, 3 March 2017

Background

- Diagrams are an important teaching means in STEM
- Accessibility is key to inclusive education
- Avoid additional hurdles for both authors and readers
- Fully automatic to reduce need for human intervention

- Technology fully functional for Chemistry diagrams
- Current work involves
  - Physics diagrams (with John Gardner and ViewPlus)
  - Statistics Graphs (with J. Godfrey and D. Fitzpatrick)

Pathways to Accessibility

- Web Accessibility (Volker)
  - Access via Screen Reader and Browser
  - Exploration with keyboard or mouse interaction
- Audio Tactile Diagrams (John)
  - Combination of embossed diagrams and audio feedback
  - Using IVEO Reader

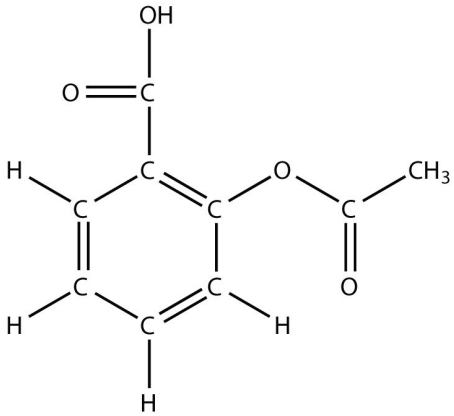
Automatic Generation of Accessible SVG

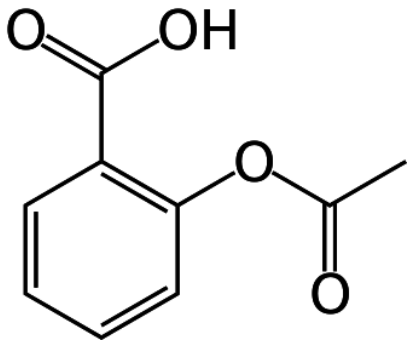
- Start with (bitmap) images of diagrams
- Understand their content
- Enrich information specifically for teaching
- Generate "content-heavy" SVG

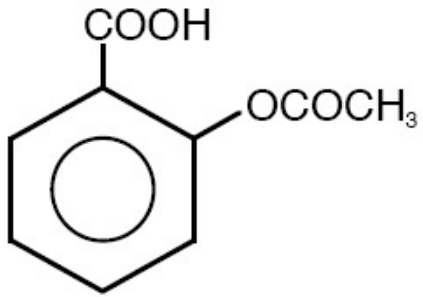
Do all this fully automatically without user interaction

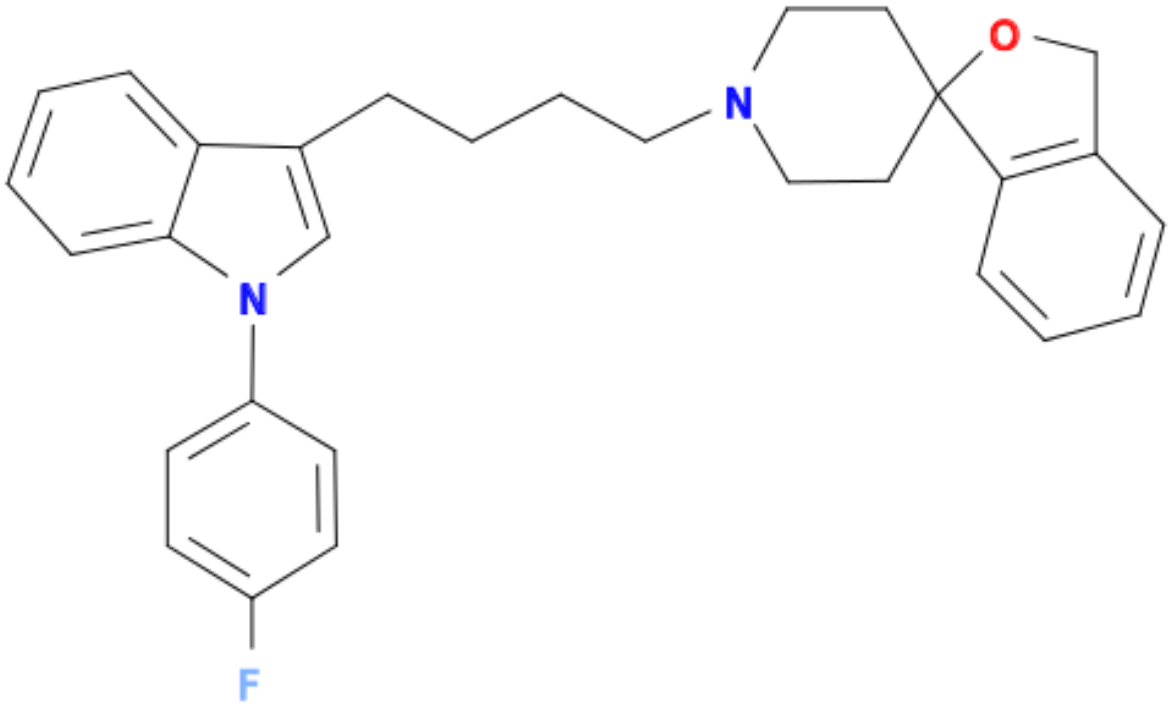
- Source independence
  - Do not rely on the benevolent, educated author
- Tool Independence
  - Do not force readers to use a specialist software tool
- Platform independence
  - Accessible with all browsers, screen readers, on all platforms

Chemistry: Molecule Diagrams

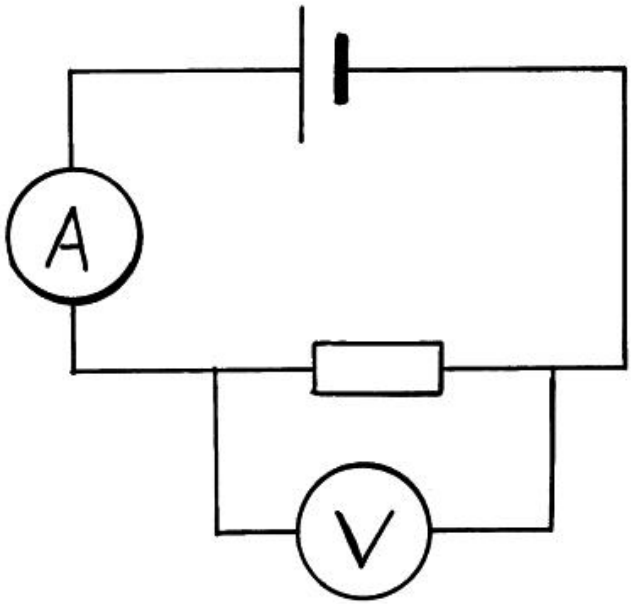


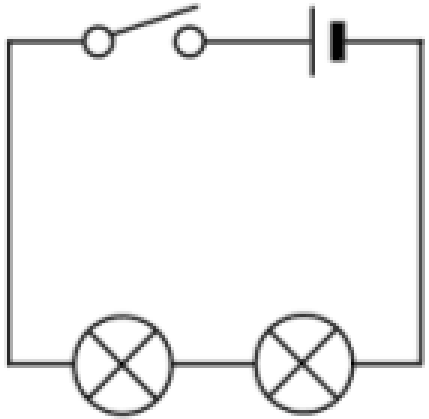


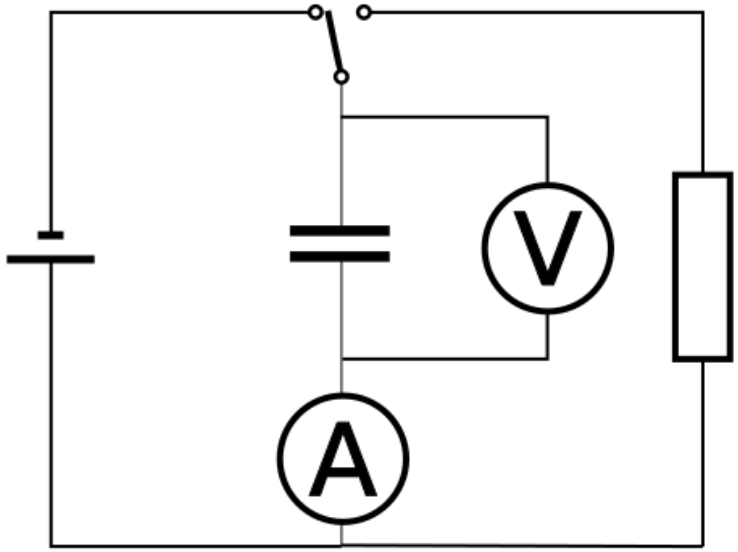




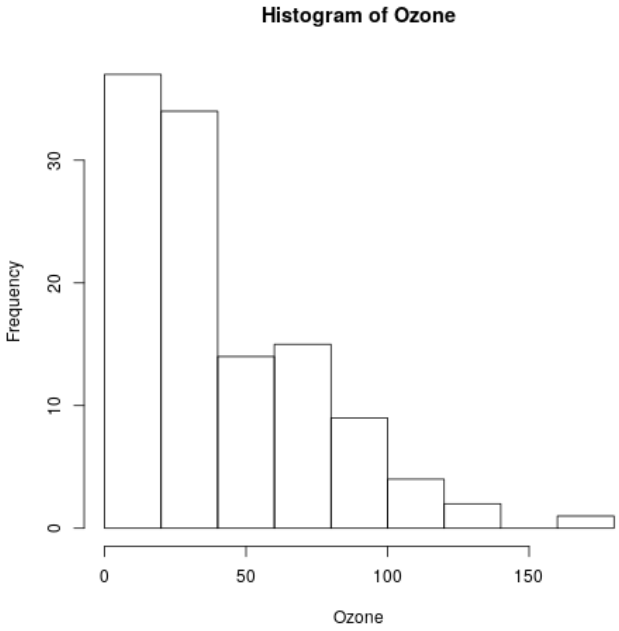
Physics: Circuit Diagrams

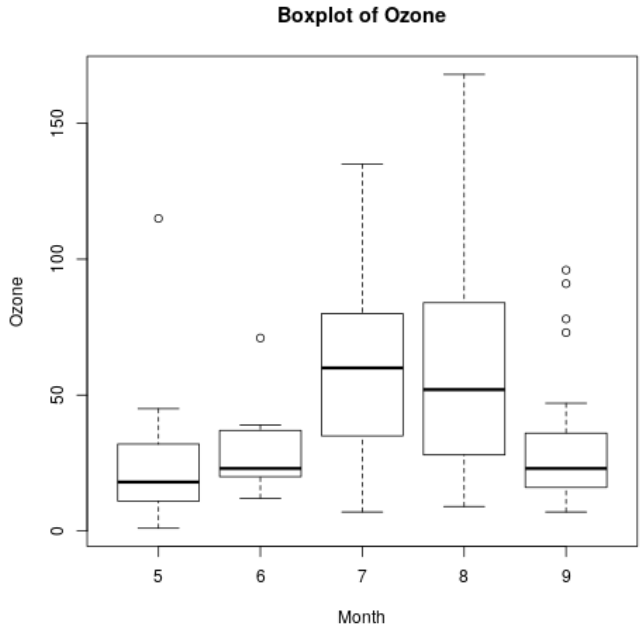


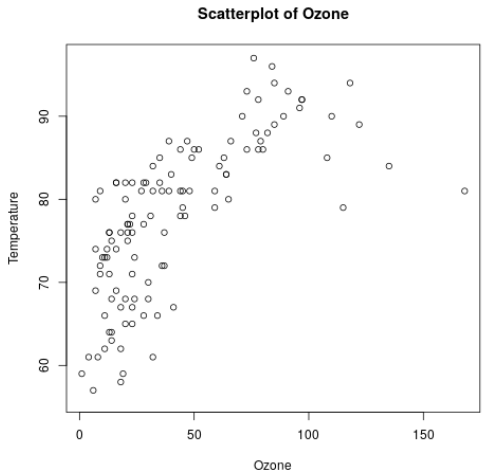


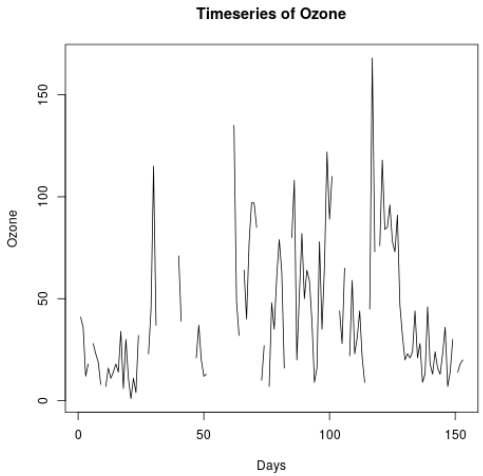


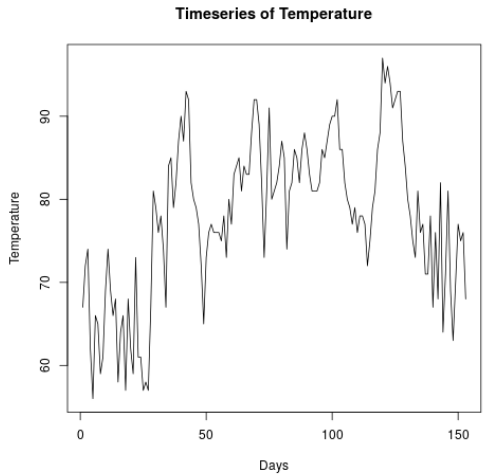
Statistics: Discrete and Continuous Data











End-to-end procedure from images to accessible diagrams

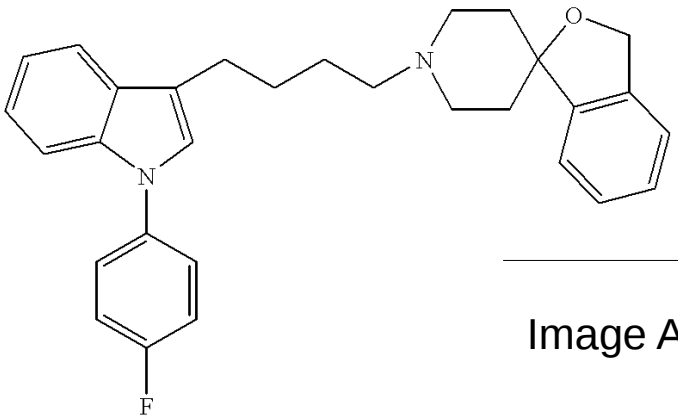
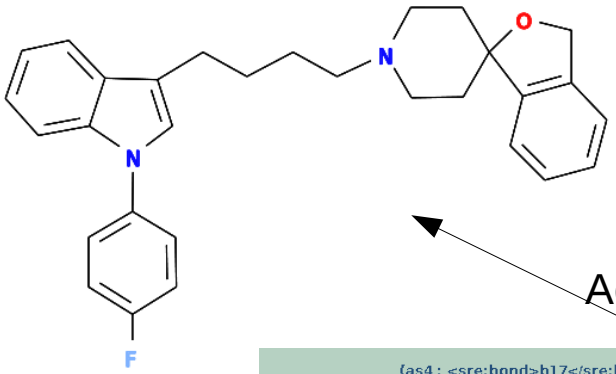


Image Analysis

```
9;3;789;482;206.732374;474.686053
chargroup;O;660;5;677;23
chargroup;N;485;57;505;75
chargroup;F;165;459;179;477
chargroup;N;162;192;182;210
line;normal;627;65;659;26;3;-1.216908
line;normal;595;122;627;66;3;-1.757429
line;normal;669;120;628;66;3;1.313009
line;normal;205;98;140;98;3;0.000000
line;normal;119;160;139;99;3;-3.003894
line;normal;670;119;732;103;3;-0.266705
line;normal;764;213;698;231;3;-0.262990
line;normal;779;150;733;103;3;1.031856
line;normal;677;126;731;112;3;-0.272006
line;normal;57;166;112;154;3;-0.231100
line;normal;172;453;172;402;3;4611686018427387904.000000
line;normal;172;390;123;362;3;0.578247
line;normal;222;362;222;304;3;4611686018427387904.000000
line;normal;123;304;171;276;3;-0.580102
line;normal;229;300;173;267;3;0.573762
line;normal;230;365;230;301;3;4611686018427387904.000000
line;normal;173;400;230;366;3;-0.574918
line;normal;115;300;171;267;3;-0.573762
line;normal;114;365;114;301;3;4611686018427387904.000000
line;normal;172;401;114;366;3;0.576680
line;normal;172;266;172;216;3;4611686018427387904.000000
```

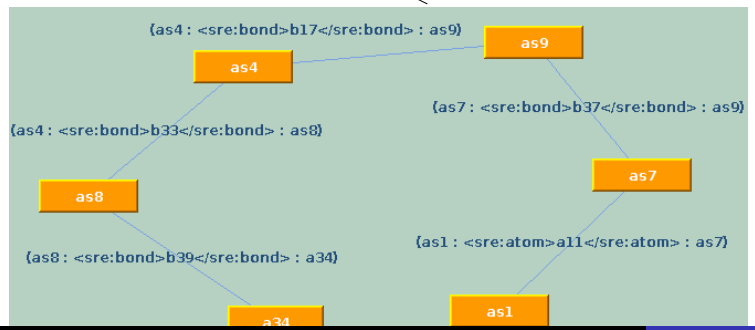
Image Recognition



Accessible Rendering

```
<?xml version="1.0"?>
<molecule id="Molfile US06358966-20020319-C00001.tif" xmlns="http://www
 <atomArray>
  <atom id="a1" elementType="C" x2="5.252700" y2="-0.976900"/>
  <atom id="a2" elementType="C" x2="-0.897300" y2="-1.526900"/>
  <atom id="a3" elementType="C" x2="-0.907300" y2="-3.576900"/>
  <atom id="a4" elementType="C" x2="0.202700" y2="-2.946900"/>
  <atom id="a5" elementType="C" x2="-0.347300" y2="-2.616900"/>
  <atom id="a6" elementType="C" x2="4.472700" y2="-1.776900"/>
  <atom id="a7" elementType="C" x2="5.702700" y2="-1.446900"/>
  <atom id="a8" elementType="C" x2="-0.027300" y2="-0.936900"/>
  <atom id="a9" elementType="C" x2="-1.757300" y2="-0.566900"/>
  <atom id="a10" elementType="C" x2="-1.157300" y2="-0.436900"/>
  <atom id="a11" elementType="C" x2="4.202700" y2="-0.586900"/>
  <atom id="a12" elementType="C" x2="-0.697300" y2="-0.916900"/>
  <atom id="a13" elementType="C" x2="5.542700" y2="-2.046900"/>
  <atom id="a14" elementType="C" x2="-0.347300" y2="-3.926900"/>
  <atom id="a15" elementType="C" x2="0.212700" y2="-3.576900"/>
  <atom id="a16" elementType="C" x2="-0.907300" y2="-2.946900"/>
```

Semantic Enrichment



Workflow

Input: A bitmap image of a molecule diagram

1. Image analysis
2. Image recognition
3. Semantic enrichment
4. Generation of annotated SVG
5. Accessible diagram via browser front-end

Modularity

Steps are modular

- and we can start/stop in any stage,
- depending what information we initially have
- or what output we want to produce

For example,

- Chemistry: from image or with chemical identifiers (e.g., from Database)
- Physics: from image only
- Statistics: from the R package using the full statistical model

Image Analysis and Recognition

Analysis: Generic procedure to extract geometric primitives

- Preprocessing: Binarisation, noise reduction...
- Connected component extraction and OCR
- Result is a set of geometric primitives:
  Character groups, lines, circles, triangles

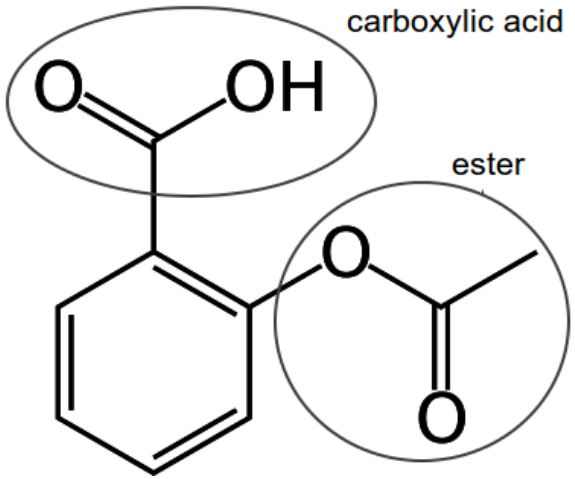
Recognition: Domain specific rewriting system

- Rule based rewriting of primitives into XML representation.
  E.g.,
- Chemistry: Molecule in Chemical Markup (CML or MOL)
- Physics: Circuit in bespoke XML

Semantic Enrichment

- Recognition phase can only generate basic information
- For teaching/learning more detail is needed
- Detect semantically interesting substructures
- Describe them and order them by importance following expert rules.

- Chemistry Example:
  - Aliphatic chains
  - Ring systems: Isolated and fused
  - Functional groups

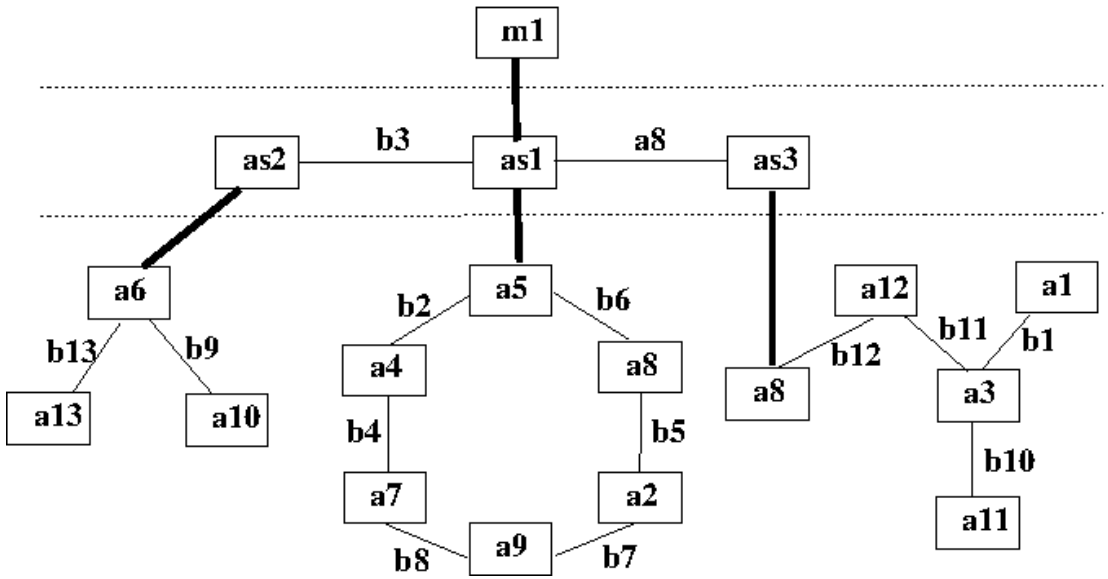


- Order blocks and atoms by chemical conventions

Abstraction and Navigation Model

- Represent diagram as layered graph
- Abstract and summaries subelements via multiple layers
- Distinguish active and passive elements for display and navigation

- Chemistry Example:



- Molecule summary
- Component summary
- Atoms and bonds

Accessibility Support

- Simple navigation model with arrow keys
- Screen Reader Support:
  - Generate speech output on different levels
  - Display of speech output using subtitling
- Low Vision/Learning Disability Support:
  - Highlighting of inspected components
  - Optional zooming and magnification of components
  - Changing contrast, colour configurations
- I18n support

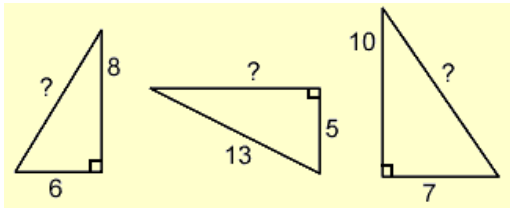
Future Support for Other Disabilities

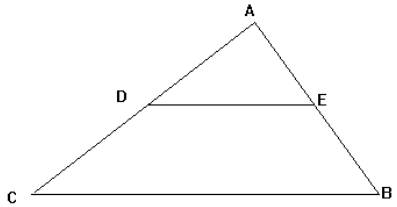
- More low-vision options
- Improved support for students with dyslexia/dysgraphia

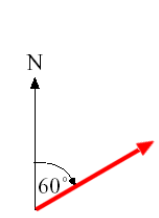
- More mouse interaction
- Automatic reading
- Synchronised shading

Future Subject Areas

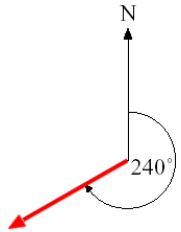
- Other STEM subjects: Maths (geometry, bearings), Biology (systems diagrams), Computer Science (flow charts)



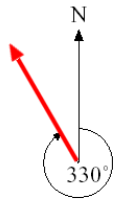




Bearing 060 °



Bearing 240 °



Bearing 330 °