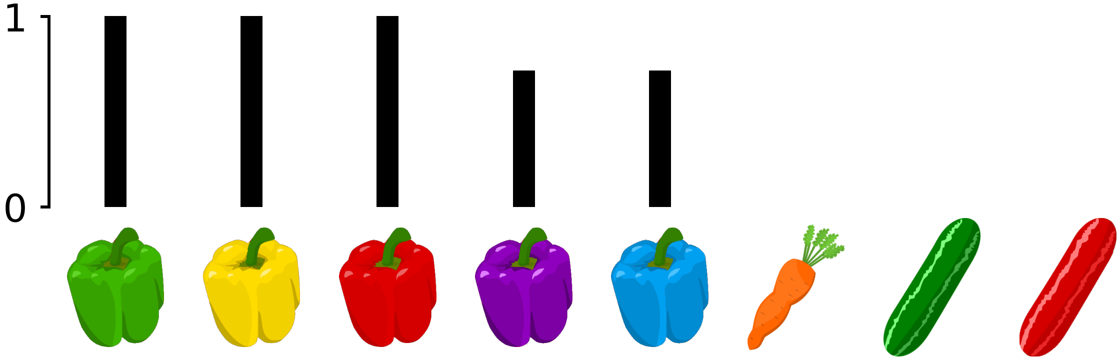
# Bernoulli Fields

## Probability distributions for vague semantics and pragmatics

Guy Emerson

# Truth-Conditional Functions



# Probabilities of Truth

- For each pixie $x$, truth-valued random variable $T_x$

# Probabilities of Truth

- For each pixie $x$, truth-valued random variable $T_x$
- $f$ : pixie $\mapsto$ probability of truth

# Probabilities of Truth

- For each pixie $x$, truth-valued random variable $T_x$
- $f$ : pixie $\mapsto$ probability of truth
  - $\mathbb{P}(T_x{=}\top) = f(x), \qquad \mathbb{P}(T_x{=}\bot) = 1 - f(x)$

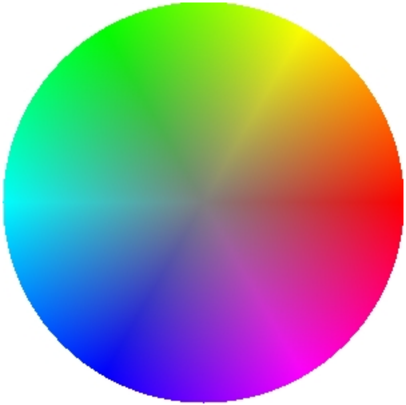
# Probabilities of Truth

- For each pixie $x$, truth-valued random variable $T_x$
- $f$ : pixie $\mapsto$ probability of truth
  - $\mathbb{P}(T_x{=}\top) = f(x)$, $\qquad \mathbb{P}(T_x{=}\bot) = 1 - f(x)$
  - $\mathbb{P}\left(T_y{=}\top\right) = f(y)$, $\qquad \mathbb{P}\left(T_y{=}\bot\right) = 1 - f(y)$

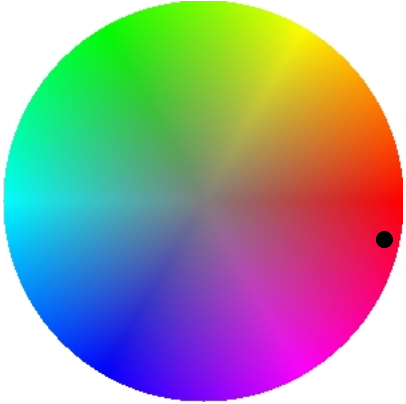
# Probabilities of Truth

- For each pixie $x$, truth-valued random variable $T_x$
- $f$ : pixie $\mapsto$ probability of truth
  - $\mathbb{P}\left(T_x{=}\top\right) = f(x)$, $\qquad \mathbb{P}\left(T_x{=}\bot\right) = 1 - f(x)$
  - $\mathbb{P}\left(T_y{=}\top\right) = f(y)$, $\qquad \mathbb{P}\left(T_y{=}\bot\right) = 1 - f(y)$
- What about $\mathbb{P}\left(T_x{=}\top, T_y{=}\top\right)$?

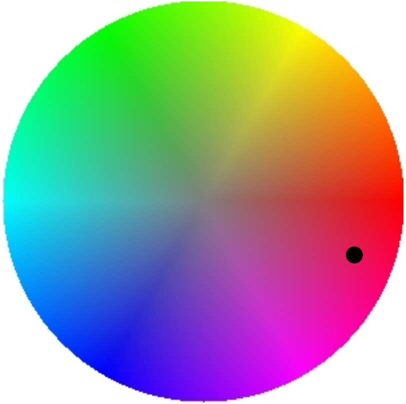
# Example: Colour Space



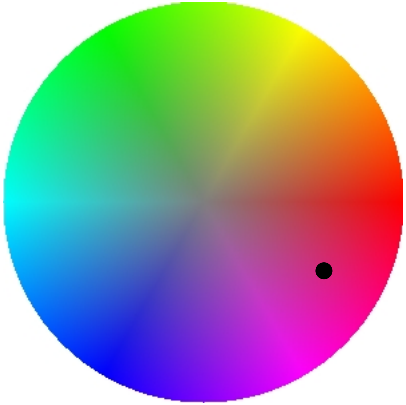
# Example: Colour Space



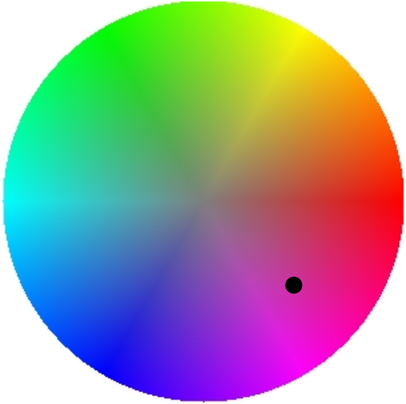
# Example: Colour Space



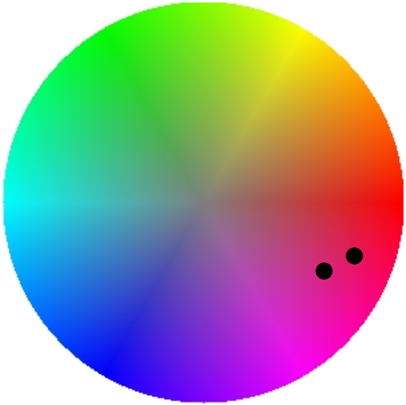
# Example: Colour Space



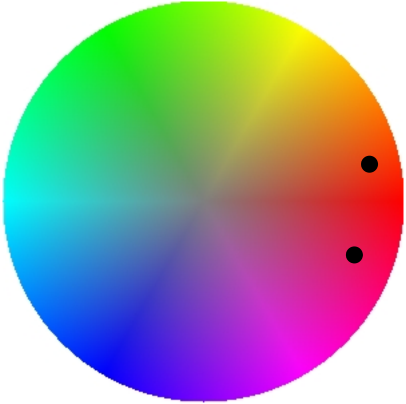
# Example: Colour Space



# Example: Colour Space



# Example: Colour Space



# Probabilistic but Coherent

- Language use is probabilistic: the extension of “red” varies from discourse to discourse

# Probabilistic but Coherent

- Language use is probabilistic: the extension of “red” varies from discourse to discourse
- Language use is coherent: the extension of “red” is self-consistent within a discourse

# Jointly Distributed Truth Values

| | $T_y = \perp$ | $T_y = \top$ |
|---|---|---|
| $T_x = \perp$ | $p_{\perp\perp}$ | $p_{\perp\top}$ |
| $T_x = \top$ | $p_{\top\perp}$ | $p_{\top\top}$ |

# Jointly Distributed Truth Values

| | $T_y = \bot$ | $T_y = \top$ |
|---|---|---|
| $T_x = \bot$ | $p_{\bot\bot}$ | $p_{\bot\top}$ |
| $T_x = \top$ | $p_{\top\bot}$ | $p_{\top\top}$ |

- 4 outcomes; probabilities sum to 1, so 3 degrees of freedom

# Jointly Distributed Truth Values

| | $T_y = \bot$ | $T_y = \top$ |
|---|---|---|
| $T_x = \bot$ | $p_{\bot\bot}$ | $p_{\bot\top}$ |
| $T_x = \top$ | $p_{\top\bot}$ | $p_{\top\top}$ |

- 4 outcomes; probabilities sum to 1, so 3 degrees of freedom

- $f(x)$ and $f(y)$ are 2 degrees of freedom

# Jointly Distributed Truth Values

- Degrees of dependence:
  - No dependence: $p_{\top\top} = f(x)f(y)$
  - Max dependence: $p_{\top\top} = \min\left\{f(x), f(y)\right\}$

# Jointly Distributed Truth Values

- Degrees of dependence:
  - No dependence: $p_{\mathrm{TT}} = f(x)f(y)$
  - Max dependence: $p_{\mathrm{TT}} = \min\{f(x), f(y)\}$
- Dependence as: covariance? correlation coefficient?

# Jointly Distributed Truth Values

- Degrees of dependence:
  - No dependence: $p_{\top\top} = f(x)f(y)$
  - Max dependence: $p_{\top\top} = \min\{f(x), f(y)\}$
- Dependence as: covariance? correlation coefficient?
  - not useful as parameters:
    range depends on $f(x)$ and $f(y)$

# Jointly Distributed Truth Values

- Degrees of dependence:
  - No dependence: $p_{\top\top} = f(x)f(y)$
  - Max dependence: $p_{\top\top} = \min\left\{f(x), f(y)\right\}$

- Dependence as **odds ratio**: $\omega = \dfrac{p_{\top\top}p_{\bot\bot}}{p_{\top\bot}p_{\bot\top}}$

# Jointly Distributed Truth Values

- Degrees of dependence:
  - No dependence: $p_{\top\top} = f(x)f(y)$
  - Max dependence: $p_{\top\top} = \min\left\{f(x), f(y)\right\}$
- Dependence as **odds ratio**: $\omega = \dfrac{p_{\top\top}p_{\perp\perp}}{p_{\top\perp}p_{\perp\top}}$
- 3 degrees of freedom: $f(x)$, $f(y)$, $\omega(x, y)$

# Local Dependence

- $f : x \mapsto$ marginal probability of truth
- $\omega : x, y \mapsto$ odds ratio (degree of dependence)
  - Coherence: if $x$ and $y$ are close, $\omega$ is high

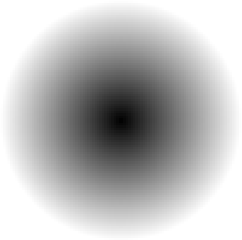
# Local Dependence

- $f : x \mapsto$ marginal probability of truth
- $\omega : x, y \mapsto$ odds ratio (degree of dependence)
  - Coherence: if $x$ and $y$ are close, $\omega$ is high
- For details, see forthcoming book chapter "Probabilistic Lexical Semantics: From Gaussian Embeddings to Bernoulli Fields"

# Distributions for Regions

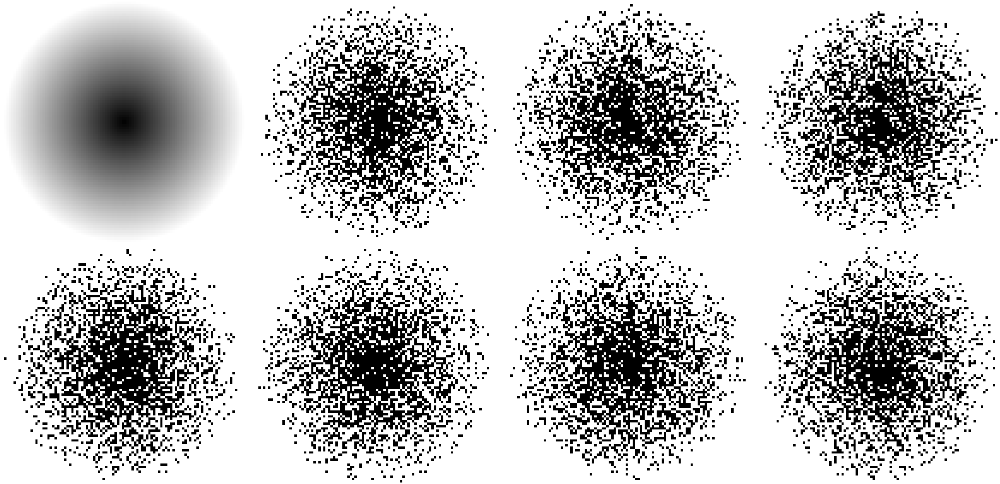
Joint distribution for truth values
$\Leftrightarrow$ Distribution for regions

# Distributions for Regions

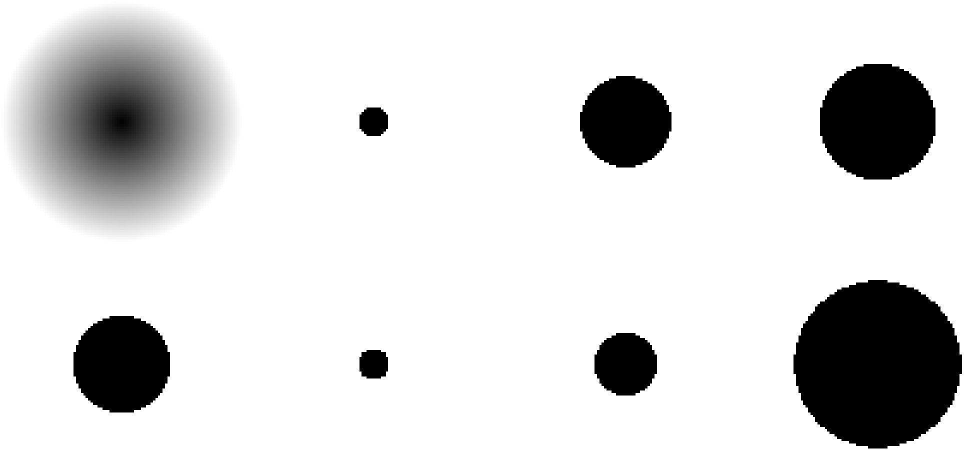


$f$, darker is higher probability of truth

# Independence



# Maximum Dependence



# Local Dependence

# Summary

- Language is probabilistic but coherent
- Can model this with $f(x)$ and $\omega(x, y)$

# Summary

- Language is probabilistic but coherent
- Can model this with $f(x)$ and $\omega(x, y)$
- Experiments on real data underway...