A Categorical Semantics of Fuzzy Concepts in Conceptual Spaces

Sean Tull

Cambridge Quantum

sean.tull@cambridgequantum.com

ACT 2021

Outline

- **‘Application’**: We formalise Gärdenfors’ **conceptual spaces**, including **fuzzy concepts** and fuzzy conceptual processes, for AI and cognitive science.

- **‘Category theory’:** We introduce a new Markov category via the category $\mathbf{LCon}$ of convex spaces and **log-concave channels**.

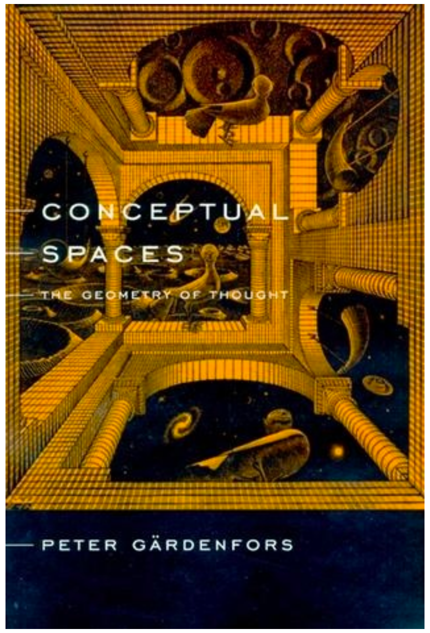
Conceptual Spaces

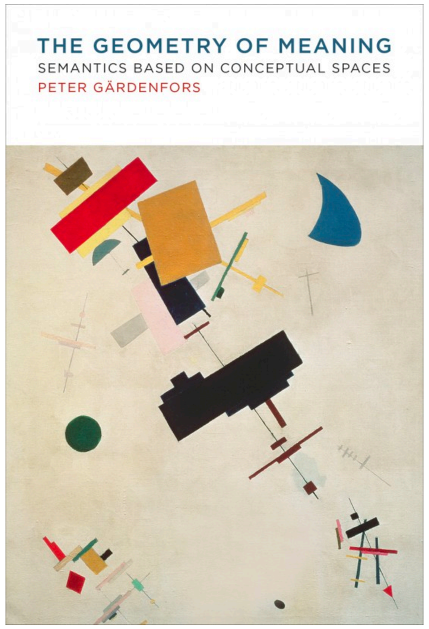
Conceptual Spaces

Framework for human and artificial cognition due to Gärdenfors.

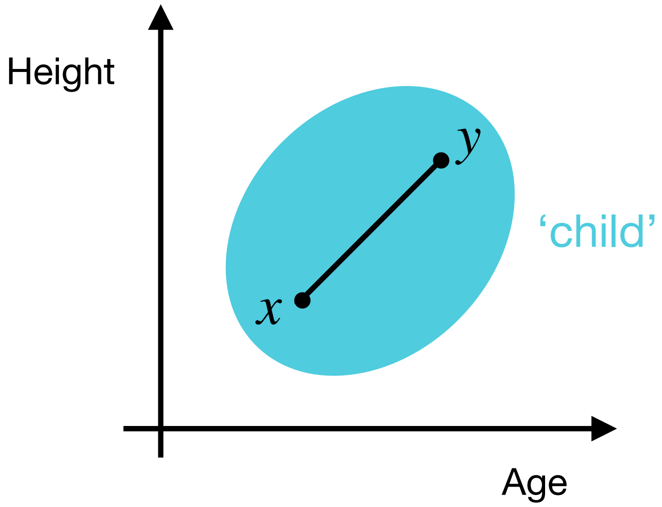
Cognitive spaces composed of **domains** e.g. colours, sounds …

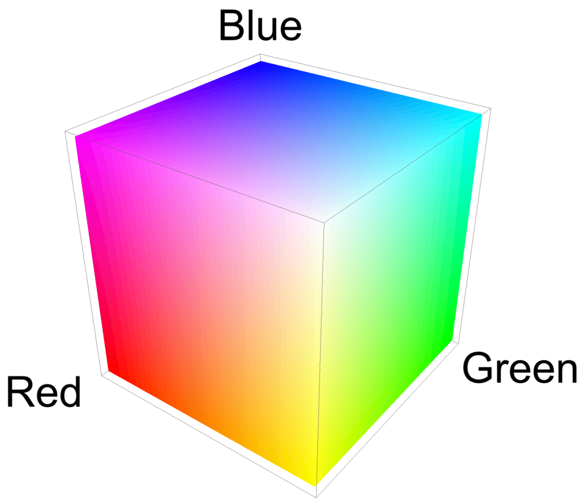
Concepts described geometrically, as regions of the space.

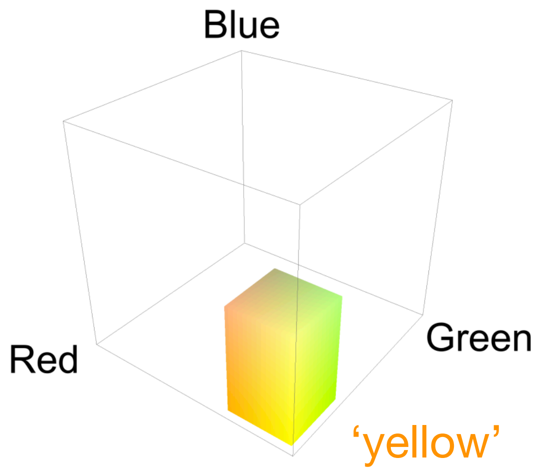




Concepts = **convex subsets**





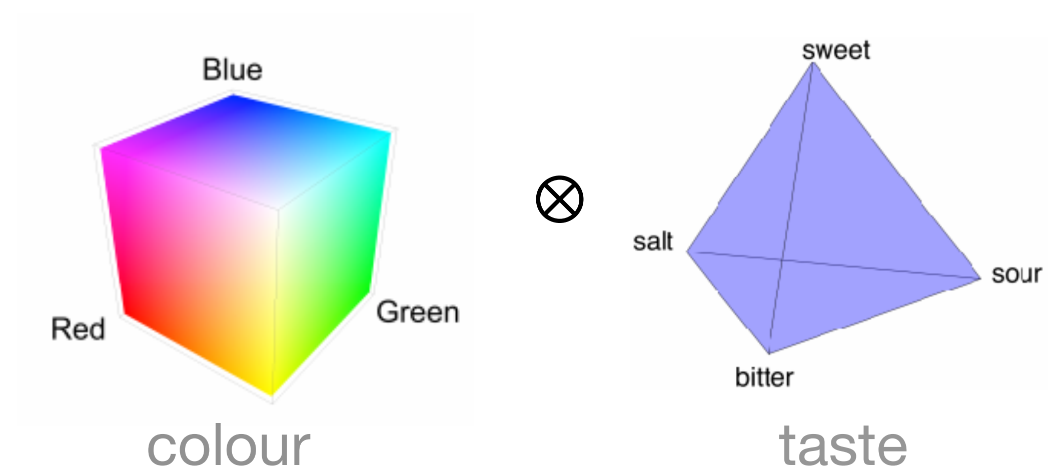


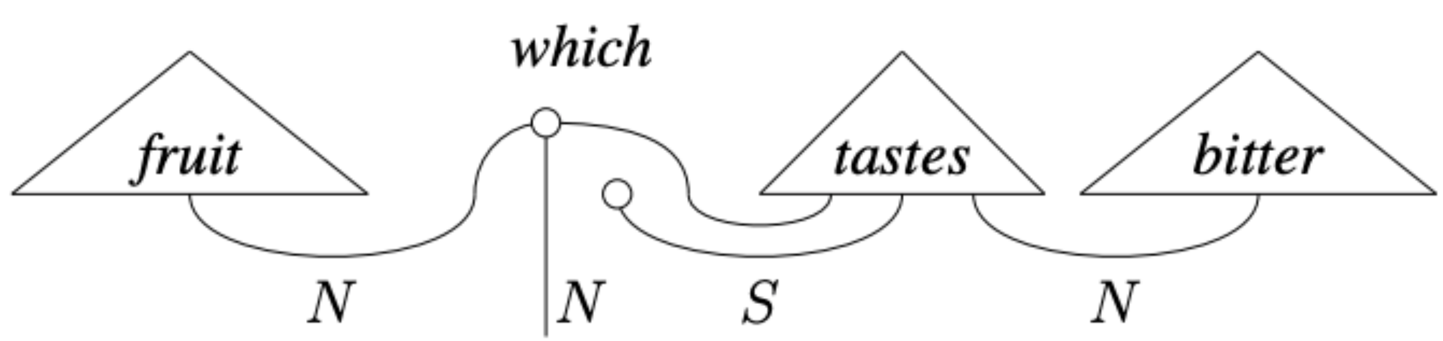
Images: (Bolt et al, 2016).

Related work

Bolt, Coecke, Genovese, Lewis, Marsden and Piedeleu, *Interacting Conceptual Spaces* (2016): formalisation via the compact category $\mathbf{ConvRel}$ of convex relations.

$N =$





See also: Vincent Wang. *Concept Functionals* (SEMSPACE 2019).

Convex Spaces

A **convex space** X is an algebra for the finite distribution monad:

$$\sum_{i=1}^{n} p_i x_i \in X \qquad\qquad (x_i \in X, \sum_i p_i = 1)$$

which is also a **measurable space**, with σ-algebra of measurable subsets

$$\Sigma_X \subseteq \mathbb{P}(X)\,.$$

A **crisp concept** of X is a measurable subset $C \subseteq X$ which is **convex:**

$$x_1, \ldots, x_n \in C \implies \sum p_i x_i \in C\,.$$

Examples of Convex Spaces

- $X = \mathbb{R}^n$ with Borel or Lebesgue σ-algebras.
- Normed space $(X, \| - \|)$, with Borel σ-algebra.
- Any convex measurable subset C of a convex space.
- Join semi-lattice L with $px + (1-p)y := x \vee y, \Sigma_L = \mathbb{P}(L)$.
- The **product** of convex spaces $X \otimes Y = X \times Y$ with

$$\sum p_i(x_i, y_i) = (\sum p_i x_i, \sum p_i y_i)\,.$$

Fuzzy Concepts

Crisp vs Fuzzy

So far concepts were 'crisp': $x \in C$ or $x \notin C$.

Evidence suggests 'real' concepts should be **fuzzy**, given as maps:

$$C : X \to [0,1]$$

$C(x)$:= "extent to which x is an instance of C."

Fuzziness also helps learning via gradient descent in neural networks.

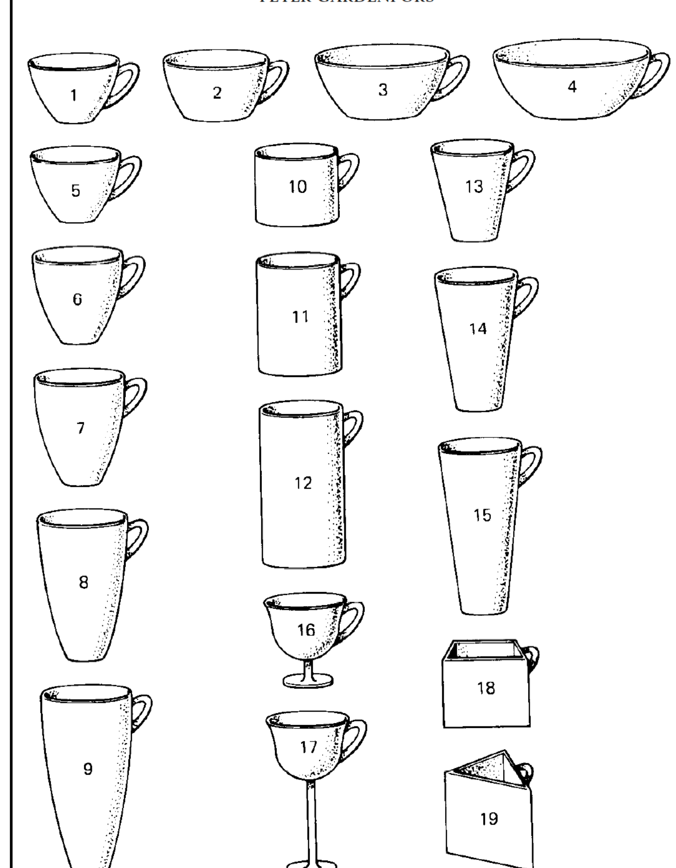

PETER GÄRDENFORS

Figure 1. Cuplike objects (From Labov 1973).

Quasi-Concavity

A natural condition is the following.

Criterion

Fuzzy concepts $C : X \to [0,1]$ should be **quasi-concave**:

$$C(px + (1-p)y) \geq \min\{C(x), C(y)\} \qquad (\forall x, y \in X, p \in [0,1])$$

Equivalently, each set $C^t := \{x \in X \mid C(x) \geq t\}$ is convex.

!

Quasi-concavity is not **compositional**: $(x, y) \mapsto C(x)D(y)$ need not be quasi-concave, even if C, D are.

e.g. $C(x) = (1-x)/2, D(y) = (y^2+1)/2$ on $[0,1]$.

Log-Concavity

Luckily, there is a class of ‘nice’ quasi-concave maps.

A function $f: X \to \mathbb{R}$ is **log-concave** (LC) if

$$f(px + (1-p)y) \geq f(x)^p f(y)^{1-p} \qquad (\forall x, y \in X, p \in [0,1])$$

Equivalently, $\log \circ f$ is concave.

LC functions and their associated measures are:

- well-studied in statistics and economics;
- well-behaved (e.g. under products, marginals, convolutions);
- a functional analogue of convex subsets (Klartag and Milman, 2005).

Geometry of log-concave functions and measures (Klartag and Milman, 2005).

Log-concavity and strong log-concavity: a review. (Saumard, Wellner 2014).

Fuzzy Concepts

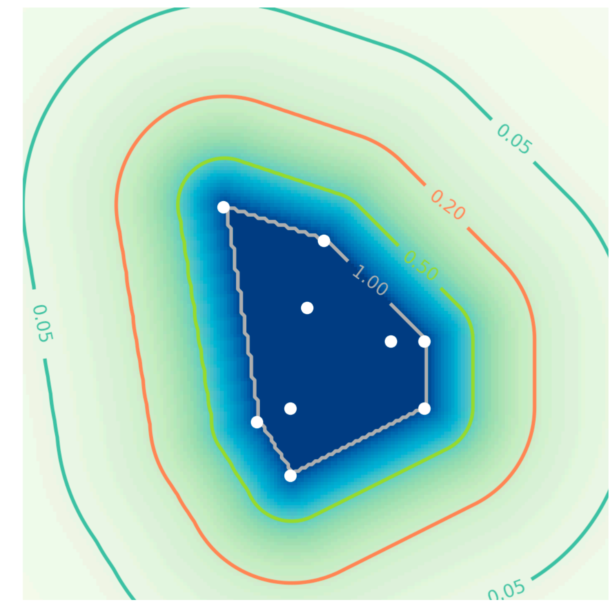
A **fuzzy concept** on a convex space X is a measurable log-concave map $C \colon X \to [0,1]$.

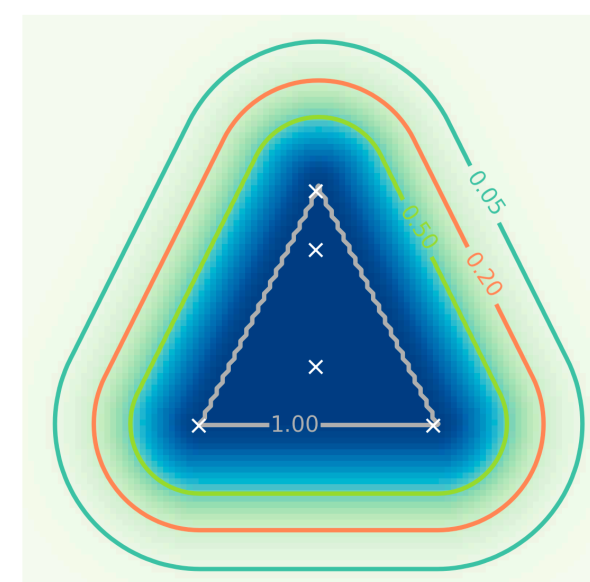
Examples

- Any crisp concept $M \subseteq X$, via its indicator $C = 1_M$.
- Measurable affine maps: $C(\Sigma p_i x_i) = \Sigma p_i C(x_i)$.
- Statistical functions on $\mathbb{R}^n$, e.g. densities of normal, exponential, logistic… distributions.
- For any crisp $P \subseteq \mathbb{R}^n$ the 'fuzzification'

$$C(x) = e^{-\frac{1}{2\sigma^2} d_H(x,P)^2}$$

where $\sigma \geq 0$, d_H is Hausdorff distance.





Log-Concavity is Canonical

Theorem

$C(X) := \{$log-concave maps$\}$ forms the largest choice of a set $C(X)$ of quasi-concave maps $X \to [0,1]$ on each convex space X such that:

- $(x, y) \mapsto C(x)D(y) \in C(X \otimes Y) \quad \forall C \in C(X), D \in C(Y);$
- $C([0,1])$ contains all affine maps.

That is, for any such choice, every function in every $C(X)$ is log-concave.

Quasi-concave + compositional $\implies$ Log-concave

Fuzzy Processes

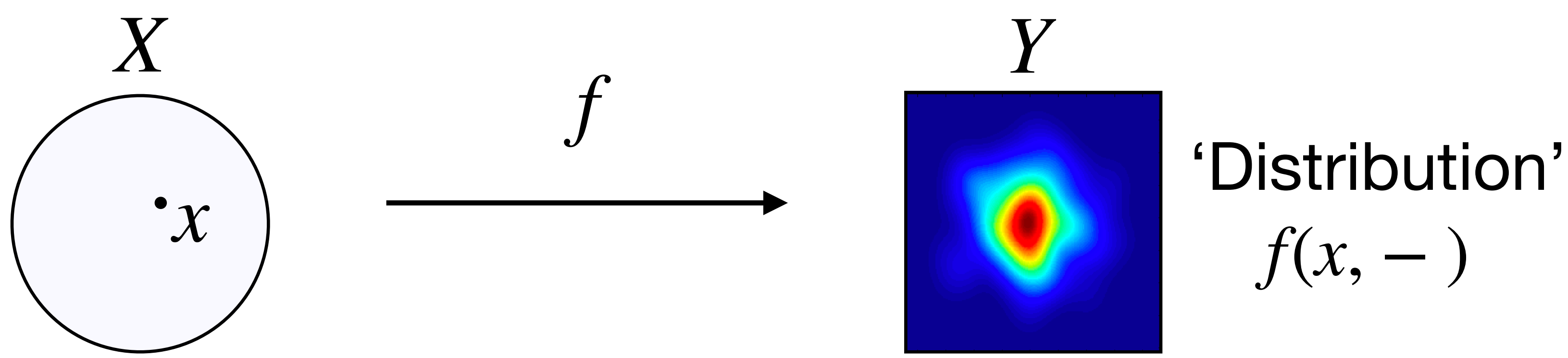
Categorical Probability

Categorical probability gives a standard notion of fuzzy map between spaces.

A probabilistic **channel**, or **Markov kernel**, $f\colon X \to Y$ is a map sending each $x \in X$ to a sub-probability measure

$$f(x, -)\colon \Sigma_Y \to [0,1]$$

over Y, in a 'measurable' way.

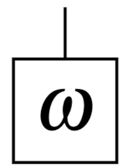


Convex spaces and channels form a symmetric monoidal category $\mathbf{Prob}$.

Abstractly, the Kleisli category of the (sub-)Giry Monad.

The Category $\mathbf{Prob}$

of probabilistic channels

States 

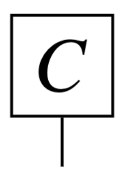
'Distributions': sub-prob measures

$\omega\colon \Sigma_X \to [0,1]$

Morphisms 

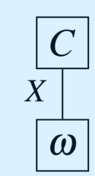
Probabilistic channels

$x \mapsto$ state $f(x, -)$ of Y

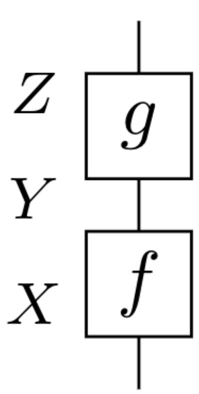
Effects 

Measurable $C\colon X \to [0,1]$

Scalars 

Probabilities e.g. 

$= \int_X C d\omega$



$:: (x, M) \mapsto \int_{y \in Y} g(y, M) df(x, y)$

$X \otimes Y = X \times Y \qquad I = \{\,\star\,\}$

Log-Concave Channels

We now generalise fuzzy concepts to conceptual channels.

We call a channel $f\colon X \to Y$ **log-concave** (or a **conceptual channel**) when

$$f(x +_p y, A +_p B) \;\geq\; f(x, A)^p f(y, B)^{1-p}$$

for crisp A, B , where $x +_p y := px + (1-p)y, \;\; A +_p B := \{pa + (1-p)b \mid a \in A, b \in B\}$.

Our main result:

Theorem

Log-concave channels form a symmetric monoidal subcategory

$$\mathbf{LCon} \hookrightarrow \mathbf{Prob}\,.$$

The Proof

The proof that $\mathbf{LCon}$ is a well-defined category is non-trivial, requiring an extension of the *Prékopa-Leindler inequality.*

Proposition

Let X be a convex space with σ-finite measures and measurable $f, g, h\colon X \to \mathbb{R}$ s.t.

$$\mu(pA + (1-p)B) \geq \nu(A)^p \omega(B)^{1-p}$$

$$f(px + (1-p)y) \geq g(x)^p h(y)^{1-p}.$$

Then

$$\left(\int_X f d\mu\right) \geq \left(\int_X g d\nu\right)^p \left(\int_X h d\omega\right)^{1-p}.$$

PL is when $X = \mathbb{R}^n, \mu = \nu = \omega$ the Lebesgue measure.

LCon is Canonical

Work in progress

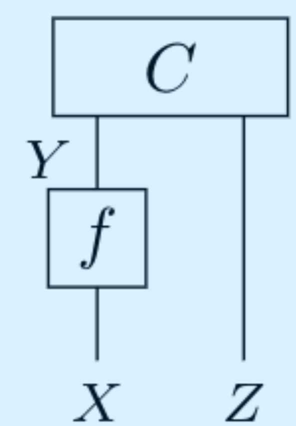
Assume Y is *well-behaved*: each set

$$\{(pa + (1-p)b, p) \mid a \in A, b \in B, p \in [0,1]\} \subseteq Y \otimes [0,1]$$

is measurable when A, B are crisp concepts (conjecture: normed spaces are).

Proposition

A channel $f\colon X \to Y$ is log-concave iff



is again a fuzzy concept, whenever C is.

It suffices to take $Z = [0,1]$ and C crisp.

$\mathbf{LCon}$ is Canonical

Work in progress

Theorem

Let $\mathbf{C}$ be a sub-SMC of $\mathbf{Prob}$ of well-behaved spaces such that

$$\{ \text{ Crisp concepts } \} \subseteq \mathbf{C}(X, I) \subseteq \{ \text{ Quasi-concave maps } \} \quad \forall X \in \mathrm{ob}\mathbf{C}$$

and $\mathbf{C}([0,1], I)$ contains all affine maps.

Then there are monoidal embeddings

$$\mathbf{C} \hookrightarrow \mathbf{LCon} \hookrightarrow \mathbf{Prob}\,.$$

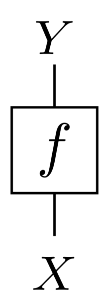
So, up to measurability considerations:

Subcategory of $\mathbf{Prob}$ + effects quasi-concave $\Longrightarrow$ Subcategory of $\mathbf{LCon}$

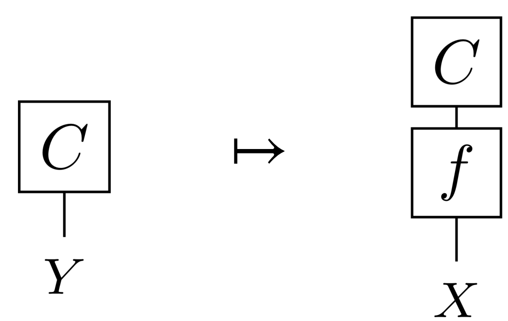
The Category $\mathbf{LCon}$

The Category $\mathbf{LCon}$

Summary so far: we propose the category $\mathbf{LCon}$ of log-concave channels $f\colon X \to Y$ between convex spaces as our category of 'fuzzy conceptual processes'.

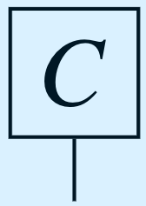
A morphism 

is a fuzzy map which preserves fuzzy concepts:

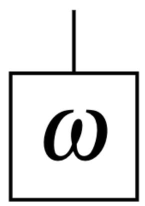


Let's meet some examples (typically in $\mathbb{R}^n$).

Morphisms in $\mathbf{LCon}$

- Effects 

 are fuzzy concepts $C\colon X \to [0,1]$.

- States 

 are sub-prob measures which are **log-concave**:

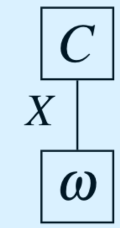
$$\omega(A +_p B) \geq \omega(A)^p \omega(B)^{1-p}$$

Log-concave measures on $\mathbb{R}^n$ include:

- Point measures δ_x;
- Uniform measures over convex regions;
- Gaussian, logistic, extreme value, Laplace…distributions;
- Lebesgue measure (Brunn-Minkowski inequality).

Morphisms in $\mathbf{LCon}$

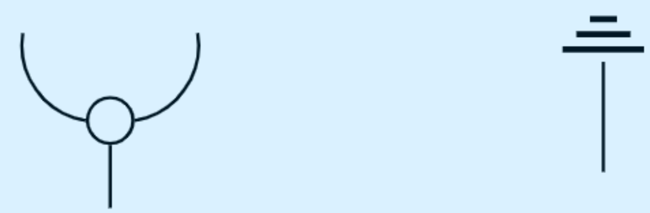
- Scalars 

 are probabilities, with 

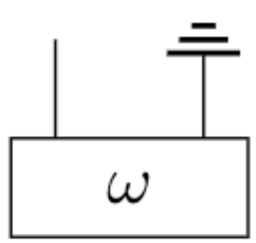
 $= \int_X C d\omega$.

- Any affine map $f \colon X \to Y$ forms a morphism via $x \mapsto \delta_{f(x)}$.

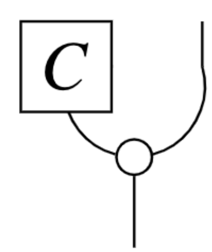
- **Copy-delete** maps, inherited from $\mathbf{Prob}$:



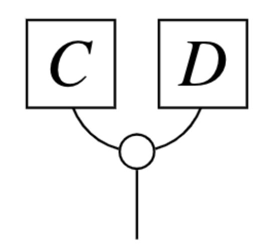
$\Longrightarrow$



Marginals



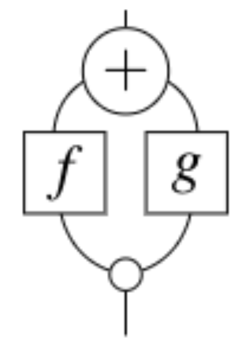
‘Update’ maps



‘Concept combinations’

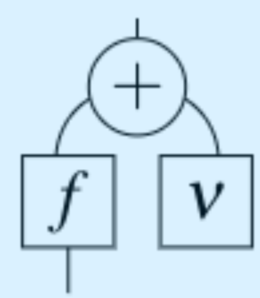
Morphisms in **LCon**

- **Convolutions** of l.c. channels into $\mathbb{R}^n$ are again l.c.:

$f \star g :=$ 

$x \mapsto$ 'sum of the random variables $f(x), g(x)$'

- For example, Gaussian noise ν added to a linear map f :



$$:: (x, A) \mapsto \nu(A - f(x))$$

This gives the subcategory of 'Gaussian probability' (Fritz 2019):

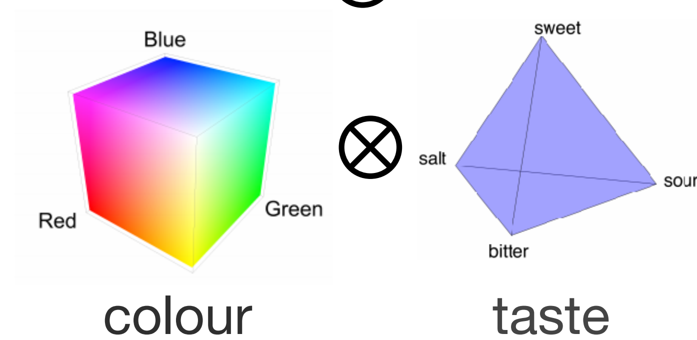
$$\mathbf{Gauss} \hookrightarrow \mathbf{LCon}$$

Toy Conceptual Reasoning

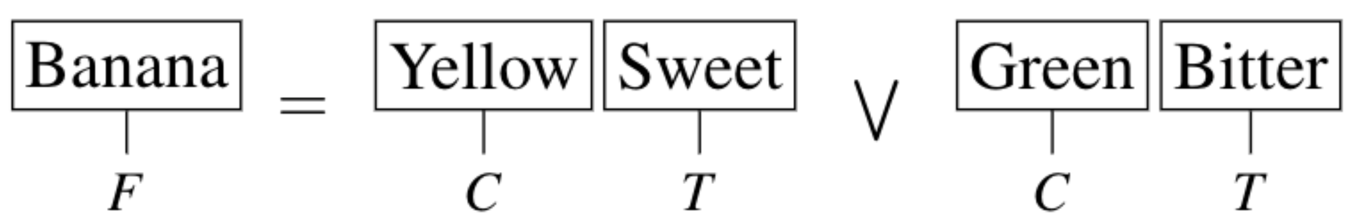
Toy Conceptual Reasoning

Following (Bolt et al.) we define a simple ‘food space’

$$F = C \otimes T$$



Can ‘learn’ a crisp concept from crisp exemplars:

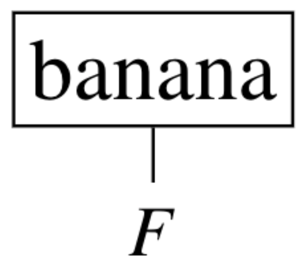


$\bigvee$ = convex closure

where $\text{Yellow} \subseteq C$ etc.

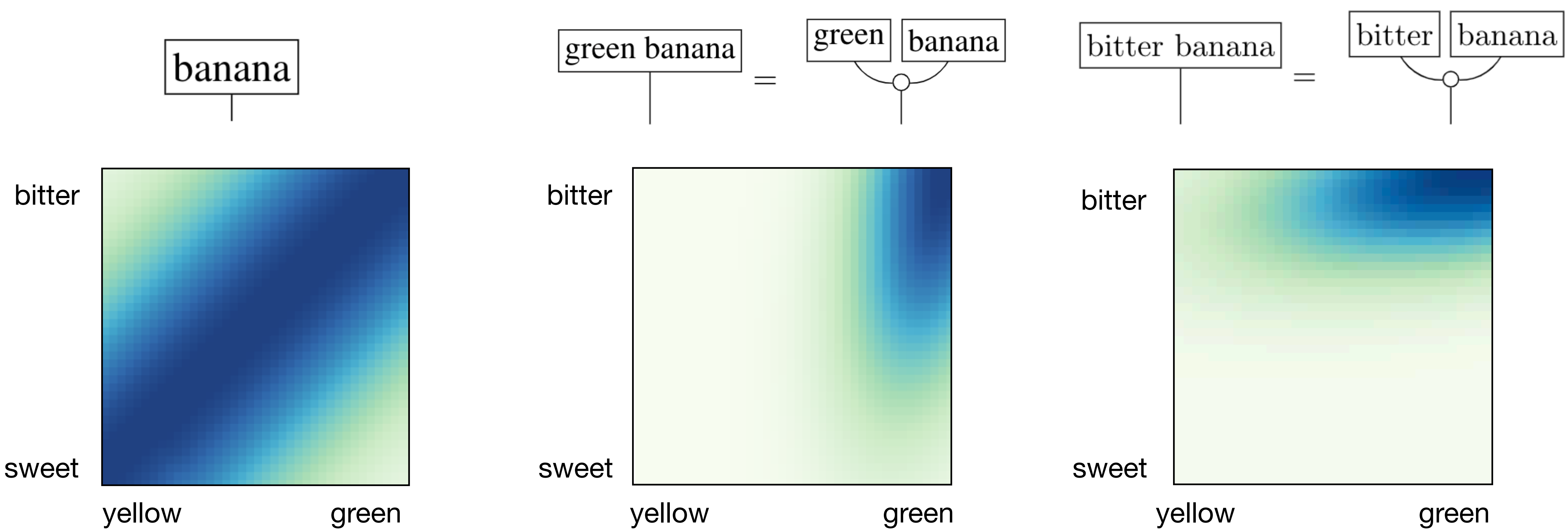
Toy Conceptual Reasoning

We can ‘fuzzify’ any of these crisp concepts e.g.



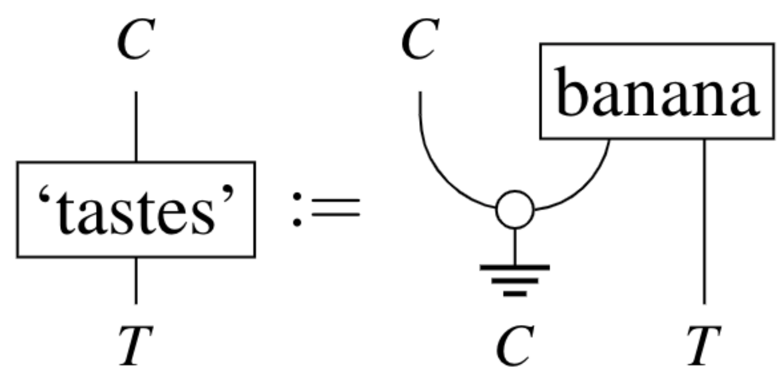
$$:: x \mapsto e^{-\frac{1}{2\sigma^2} d_H(x, \text{Banana})^2}$$

and combine them using copy/update maps:



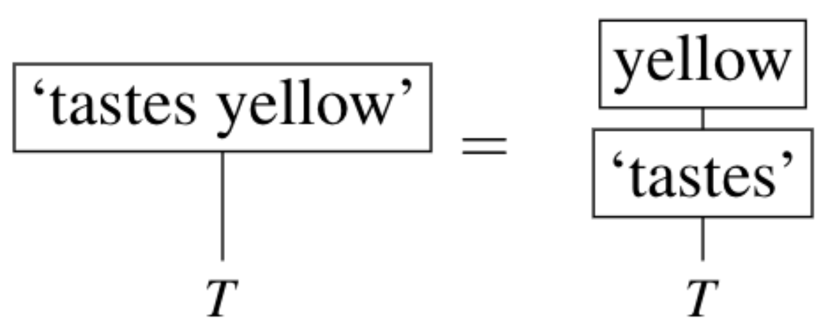
Toy Conceptual Reasoning

A simple ‘metaphor’ channel:




= uniform Lebesgue measure

This transforms colour concepts to taste concepts.



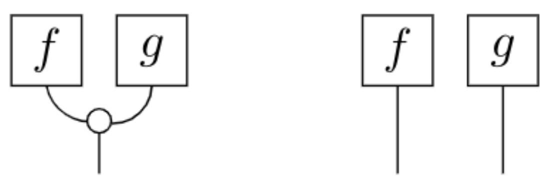
$$:: t \mapsto \frac{1}{\lambda(C)} \int_C \text{yellow}(c)\text{banana}(c,t)dc$$

Future: Explore more sophisticated conceptual ‘reasoning’ channels.

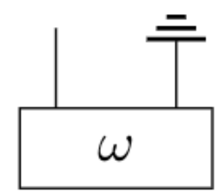
Outlook

Outlook

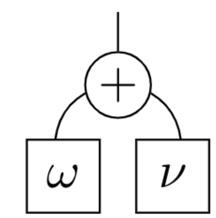
- Proposed **LCon** as a model of 'fuzzy reasoning in conceptual spaces'.
- More broadly, an interesting new Markov/copy-delete category.
- Categorical view generalises known results LC measures, e.g. closure under



Products



Marginals



Convolutions

Future work:

- Applications of fuzzy concepts in AI and NLP.
- More sophisticated channel examples e.g. 'reasoning', 'metaphor'.
- Categorical properties of **LCon**.

Thanks!