Cursus Cryptografie

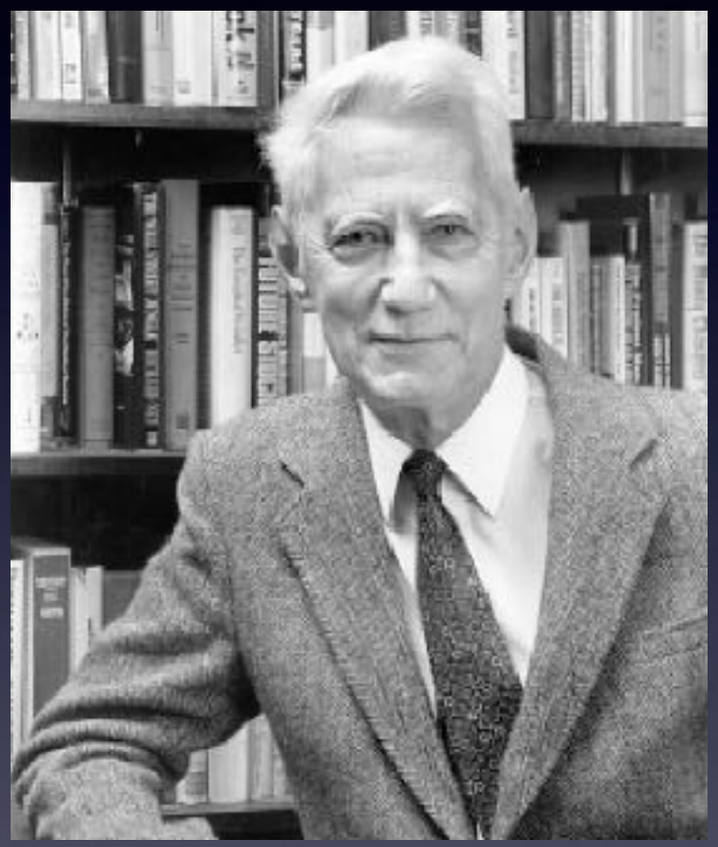
INFORMATIE

09-03-2017

Onderwerpen

- Claude Shannon
- Informatiekanaal
- Entropie
- Equivocatie
- Markov ketens
- Uniciteitsafstand
- Binair symmetrisch kanaal
- Kansberekening

Claude Shannon

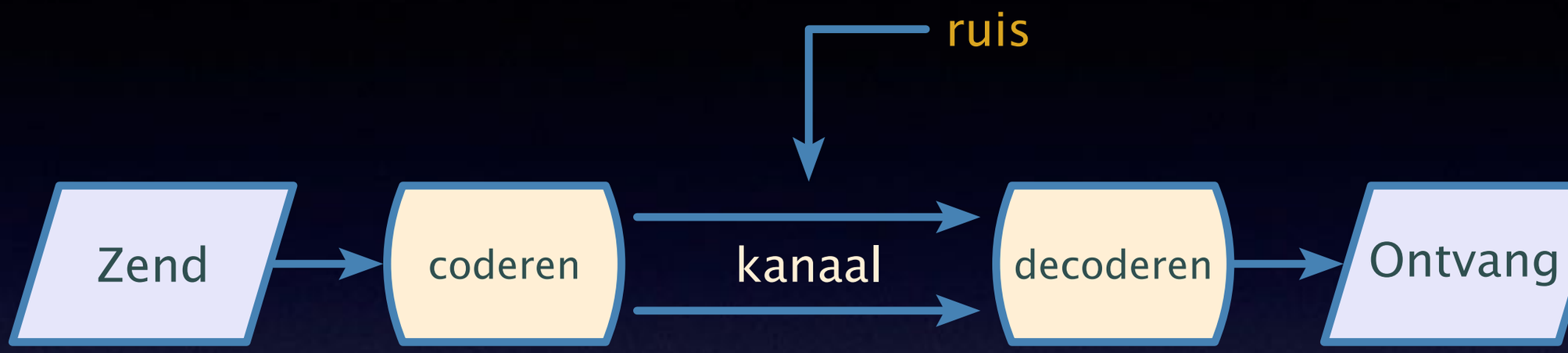


Claude Elwood Shannon (1916–2001)
AT&T Bell Telephones (1941–1972)
grondlegger van de informatietheorie

Grondslagen informatietheorie 1948
A Mathematical Theory of Communication
Bell System Technical Journal

Toepassing op cryptografie 1949
Communication Theory of Secrecy Systems
Bell System Technical Journal

Informatiekanaal



Probleemstellingen

- volledige overdracht ondanks ruis
- zo compact mogelijk coderen
- afluisteren verhinderen

Entropie

Informatiewaarde stoelt op kansbegrip

p_i = kans op symbool $s_i \in \{S\}$

$I(s_i)$ = informatiewaarde symbool s_i

$|S|$ = aantal symbolen in $\{S\}$

$$I(s_i) = \log_2 (1/p_i) = - \log_2 p_i$$

entropie van $\{S\}$ is gewogen som $= - \sum p_i \log_2 p_i$

grenswaarden $0 \leq H(S) \leq \log_2 |S|$

analogie met entropie in thermodynamica = warmteleer

Entropie Engels

	p	- $\log_2$ p	- p $\log_2$ p		p	- $\log_2$ p	- p $\log_2$ p
A	7.4%	3.763	.277	N	8.0%	3.653	.290
B	1.0%	6.682	.065	O	7.5%	3.732	.281
C	3.1%	5.027	.154	P	2.7%	5.227	.140
D	4.2%	4.558	.193	Q	0.4%	8.158	.029
E	13.0%	2.944	.383	R	7.6%	3.722	.282
F	2.8%	5.142	.146	S	6.1%	4.031	.247
G	1.6%	5.932	.097	T	9.2%	3.444	.316
H	3.4%	4.883	.165	U	2.6%	5.265	.137
I	7.4%	3.766	.277	V	1.5%	6.028	.092
J	0.2%	9.252	.015	W	1.6%	6.002	.094
K	0.3%	8.400	.025	X	0.5%	7.758	.036
L	3.6%	4.799	.174	Y	1.9%	5.692	.110
M	2.5%	5.337	.132	Z	0.1%	9.995	.010
H: 26-letters A—Z = 4.700				Engels = 4.167 bits/letter			

Equivocatie

simultane (i, j) *is afhankelijk* i *met kans* $p_{ij} = p_i p_{j|i}$

entropie $H(I, J) = -\sum_{i,j}^{|S|} p_{ij} \log p_{ij}$

$$H(I, J) = \underbrace{-\textstyle\sum_i p_i \log p_i}_{H(I)} \quad \underbrace{-\textstyle\sum_i p_i \sum_j p_{j|i} \log p_{j|i}}_{equivocatie = H(J|I)}$$

i en j onafhankelijk

$$p_{ij} = p_i p_j \quad \rightarrow \quad H(I, J) = H(I) + H(J)$$

grenswaarden

$$0 \leq H(J|I) \leq H(J) \leq \log_2 |S|$$

Zender-Ontvanger

equivocatie = kanaalverlies, noise entropy

$$H(J|I) = -\sum_i p_i \sum_j p_{j|i} \log p_{j|i}$$

perfecte communicatie

$$p_{j|i} = \begin{cases} 1 \text{ als } i = j \\ 0 \text{ als } i \neq j \end{cases} \quad \rightarrow \quad \begin{cases} H(J|I) = 0 & \text{equivocatie nul} \\ H(I,J) = H(I) & \text{dus geen verlies} \end{cases}$$

imperfecte communicatie

$$0 < p_{j|i} < 1 \;\rightarrow\; H(J|I) > 0 \;\rightarrow\; H(I,J) = H(I) + H(J|I) > H(I)$$

Wederzijdse informatie

definitie wederzijdse informatie I = *zender,* J = *cryptoanalist*

$$W(I,J) = H(J) - H(J|I)$$

perfecte waarneming → bekende sleutel

$$p_{j|i} = \begin{cases} 1 \text{ als } j = i \\ 0 \text{ als } j \neq i \end{cases} \quad \rightarrow \quad H(J|I) = 0 \quad \rightarrow \quad W(I,J) = H(J) = H(I)$$

imperfecte waarneming → cryptoanalyse

$$0 < p_{j|i} < 1 \quad \rightarrow \quad H(J|I) > 0 \quad \rightarrow \quad 0 < W(I,J) < H(J)$$

geen waarneming → onbreekbaar systeem

$$p_{j|i} = p_j \quad \rightarrow \quad H(J|I) = H(J) \quad \rightarrow \quad W(I,J) = 0$$

Markov keten

markov keten lengte 1:

kans op Qx is $p(x|Q) = \begin{cases} x = U & 1 \\ x \neq U & 0 \end{cases}$

markov keten lengte 2:

kans op QUx is $p(x|QU) = \begin{cases} x = A & 0.37 \\ x \in E, I & 0.30 \\ x = O & 0.03 \\ x \notin A, E, I, O & 0.00 \end{cases}$

markov keten lengte n:

kans op $x_1 x_2 \ldots x_n x_{n+1} \;=\; p(x_{n+1}|x_1 \ldots x_n)$

Entropie Markov keten

berekening entropie markov keten

$$p_k = \sum_{i\ldots j} p_{i\ldots j}\, p_{k|i\ldots j} \qquad p_{i\ldots k} = p_{i\ldots j} p_k$$

$$H(I\ldots JK) = -\sum_{i\ldots k} p_{i\ldots k} \log p_{i\ldots k}$$

$$H(I\ldots JK) = H(I\ldots J) + H(K|I\ldots J)$$

bepaal kansen uit lettertellingen

Entropie taal

entropie bij bericht van L letters

$$H_{L-letters} = H_1(S_1) + H_2(S_2|S_1) + H_3(S_3|S_1S_2) + \ldots$$
$$\geq L \cdot H_L(S_L|S_1 \ldots S_{L-1})$$

entropie per letter bij bericht van lengte 0, 1, 2,...

$$H_0 = 4.7 \quad (= \log_2 26)$$
$$H_1 = 4.2$$
$$H_2 = 3.6$$
$$H_\infty = 1.2\text{-}1.5$$

H_∞ = *rate of the language*

$D = H_0 - H_\infty$ = *redundantie*

Entropie Caesar

L = 2		L = 4		L = 8	
AB	0.044	ABIG	0.139	ABIGPMZM	0.000
BC	0.000	BCJH	0.000	BCJHQNAN	0.000
CD	0.002	CDKI	0.008	CDKIROBO	0.000
DE	0.142	DELJ	0.041	DELJSPCP	0.000
EF	0.088	EFMK	0.000	EFMKTQDQ	0.000
…	…	…	…	…	…
RS	0.081	RSZX	0.000	RSZXGDQD	0.000
ST	0.254	STAY	0.658	STAYHERE	0.998
TU	0.046	TUBZ	0.000	TUBZIFSZ	0.000
UV	0.007	UVCA	0.061	UVCAJGTG	0.000
…	…	…	…	…	…
ZA	0.002	ZAHF	0.003	ZAHFOLYL	0.000
$H_2 = 3.182$		$H_4 = 1.719$		$H_8 = 0.022$	

Key appearance

known-plaintext $W(K,(C,M)) = H(K) - H(K|C,M)$ *key appearance*

$$H(M,C,K) = H(C,K) + H(M|C,K)$$

$$H(M,C,K) = H(C,M) + H(K|C,M)$$

$$H(M|C,K) = 0$$

$$H(C,K) = H(C) + H(K|C)$$

$$H(C,M) = H(C) + H(M|C)$$

$H(K|C,M) = H(K|C) - H(M|C)$ *key appearance ciphertext-only*

Toont aan: known-plaintext minder onzeker dan ciphertext-only

ciphertext-only $W(C,M) = H(M) - H(M|C) \geq H(M) - H(K)$

Uniciteitsafstand

ciphertext-only: sleutel bekend? als $H(K|C) \approx 0$

$$H(K|C) = H(C,K) - H(C)$$

$$M \overset{K}{\leftrightarrow} C \rightarrow H(C,K) = H(M,K)$$

kies K *onafhankelijk* $M \rightarrow H(M,K) = H(M) + H(K)$

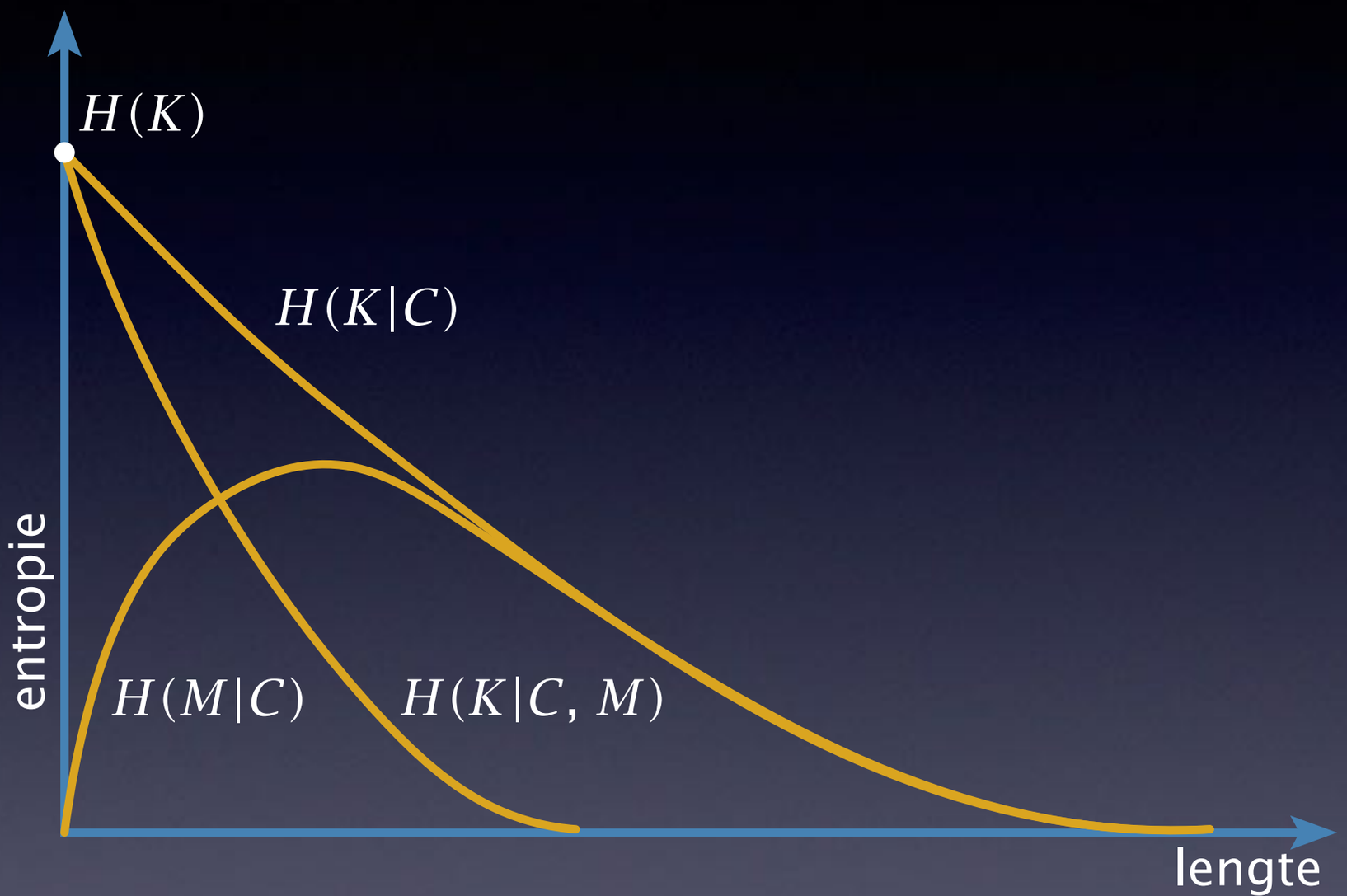
$$\rightarrow H(K|C) = H(M) + H(K) - H(C)$$

Markov model $\rightarrow H(M) \geq L \cdot H_\infty$ *en* $H(C) \leq L \cdot H_0$

$$\rightarrow H(K|C) \geq H(K) + L \cdot H_\infty - L \cdot H_0$$

$$H(K|C) \approx 0 \quad \rightarrow \quad L \approx \frac{H(K)}{H_0 - H_\infty} \quad L = \textit{uniciteitsafstand}$$

Uniciteitsafstand



H(K) = sleutelentropie

ciphertext-only H(K I C) > H(K I (C,M)) known-plaintext

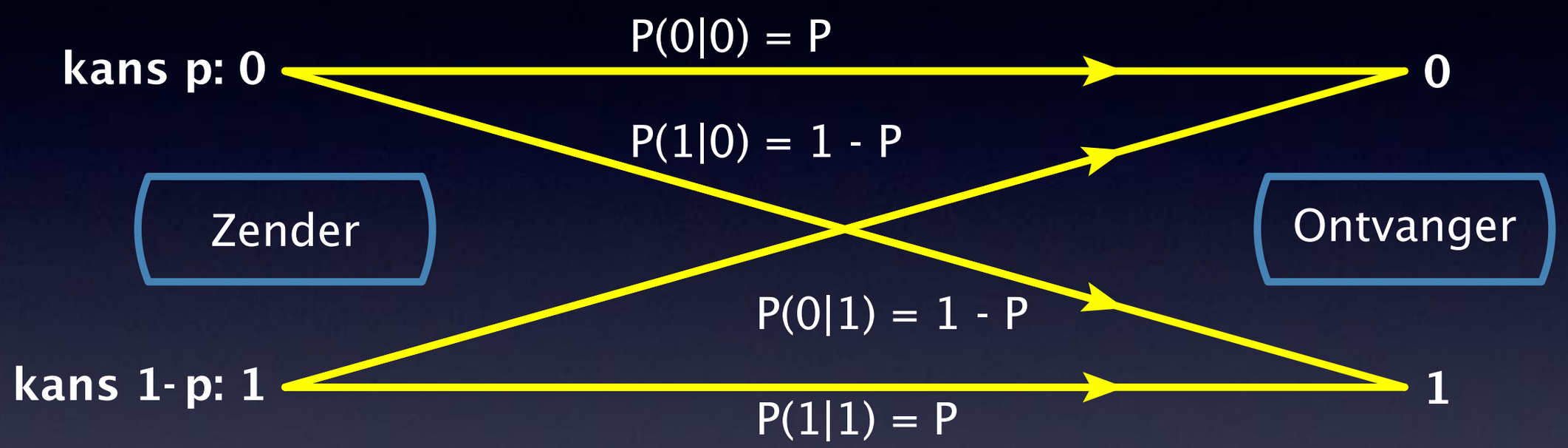
UD voorbeeld

Monoalfabeet
$L \geq H_K/(H_0-H_\infty)$ =
$\log_2 26!/(4.7-1.2) \approx$
25 letters

Transpositie			
$L \geq H_K/(H_1-H_\infty)$			
n-gram		25.10^{18}	$grille4^{16}$
n	H_n	L=20	L=8x8
1	4.2	∞	∞
2	3.6	111	55
3	3.2	65	32
7	2.8	45	22
∞	1.2	20	10

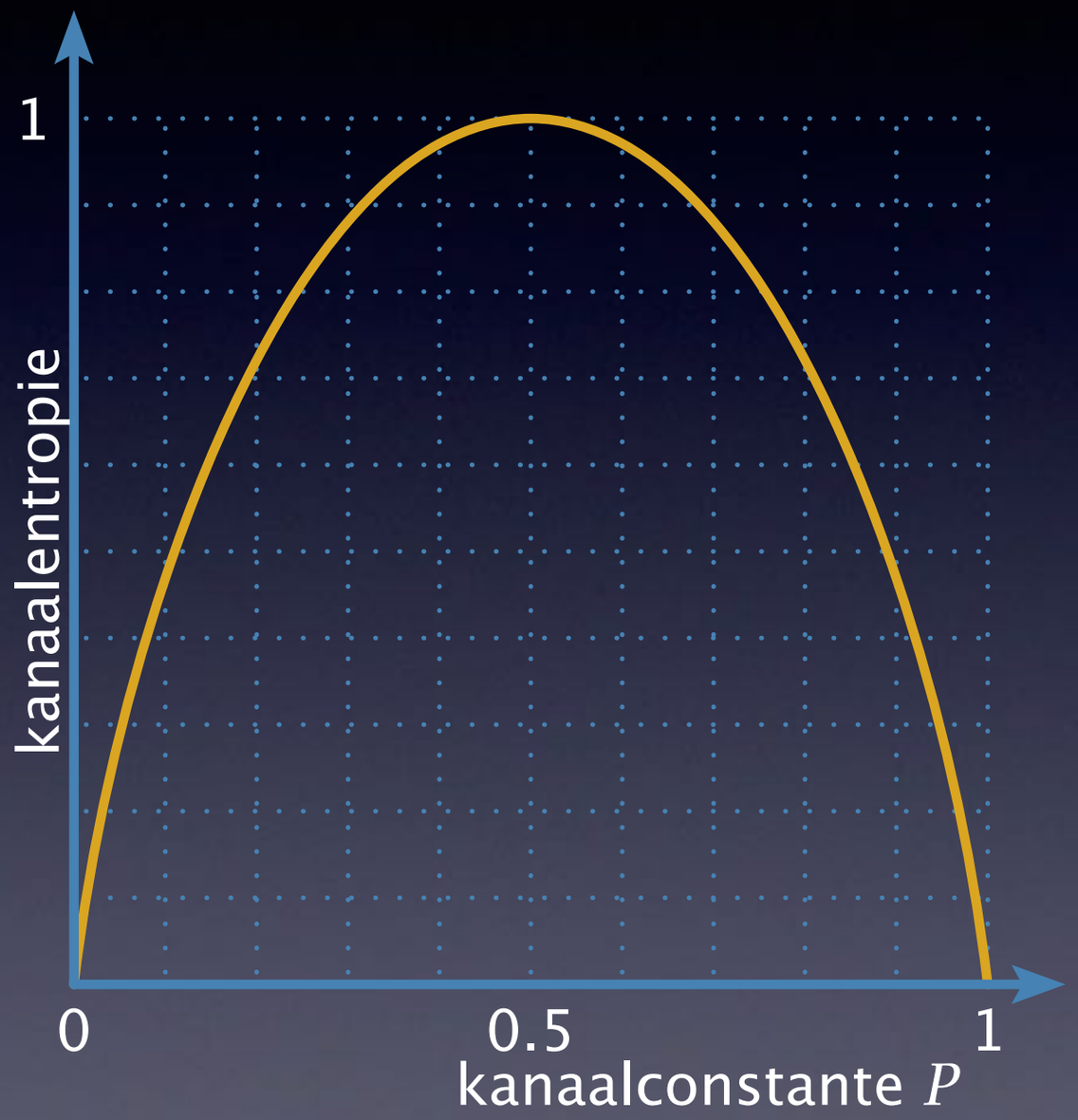
Deze *Shannon*-entropie is iets te rooskleurig
andere definities o.a. min-entropie, Renyi-entropie

Binair symmetrisch kanaal



$$P \cdot \vec{p} = \begin{pmatrix} P & 1-P \\ 1-P & P \end{pmatrix} \begin{pmatrix} p \\ 1-p \end{pmatrix} = \begin{pmatrix} 1-p-P+2pP \\ p+P-2pP \end{pmatrix}$$

Kanaalentropie



$$
\begin{aligned}
H(ontvanger|zender) = \\
&- p_0 p_{0|0} \log_2 p_{0|0} \\
&- p_0 p_{1|0} \log_2 p_{1|0} \\
&- p_1 p_{0|1} \log_2 p_{0|1} \\
&- p_1 p_{1|1} \log_2 p_{1|1} \\
= -P \log_2 P &- (1-P) \log_2 (1-P)
\end{aligned}
$$

UD kansberekening

kans op een zeker aantal oplossingen

$$\underline{M} = \#[k \in \{K\}: D(C,k) \in \{M\}] \quad \text{met} \quad \underline{M} \in \{1, \ldots, |K|\}$$

kans op goede bij meerdere oplossingen

$$\underline{M} = m \quad \rightarrow \quad p_m = \frac{1}{m}$$

kans op goede bij één of meer oplossingen

$$P = \sum_{m=1}^{|K|} \frac{1}{m} p(\underline{M} = m)$$

UD kansberekening

kans op een oplossing

$$p_k = \begin{cases} 1 & \text{goede sleutel k} \\ 2^{-L.D} & \text{willekeurige sleutel k} \end{cases}$$

schakel echte oplossing uit

$$P(\underline{M'}) \text{ met } \underline{M}' = \underline{M} - 1$$

kans bij binomiale verdeling

$$p(\underline{M}' = m') = \binom{|K| - 1}{m'} p_k^{m'} (1 - p_k)^{|K|-1-m'}$$

UD kansberekening

voeg echte oplossing weer toe

$$m' = m - 1 \quad \rightarrow \quad P = \sum_{m=1}^{|K|} \frac{1}{m} \binom{|K|-1}{m-1} p_k^{m-1} (1-p_k)^{|K|-m}$$

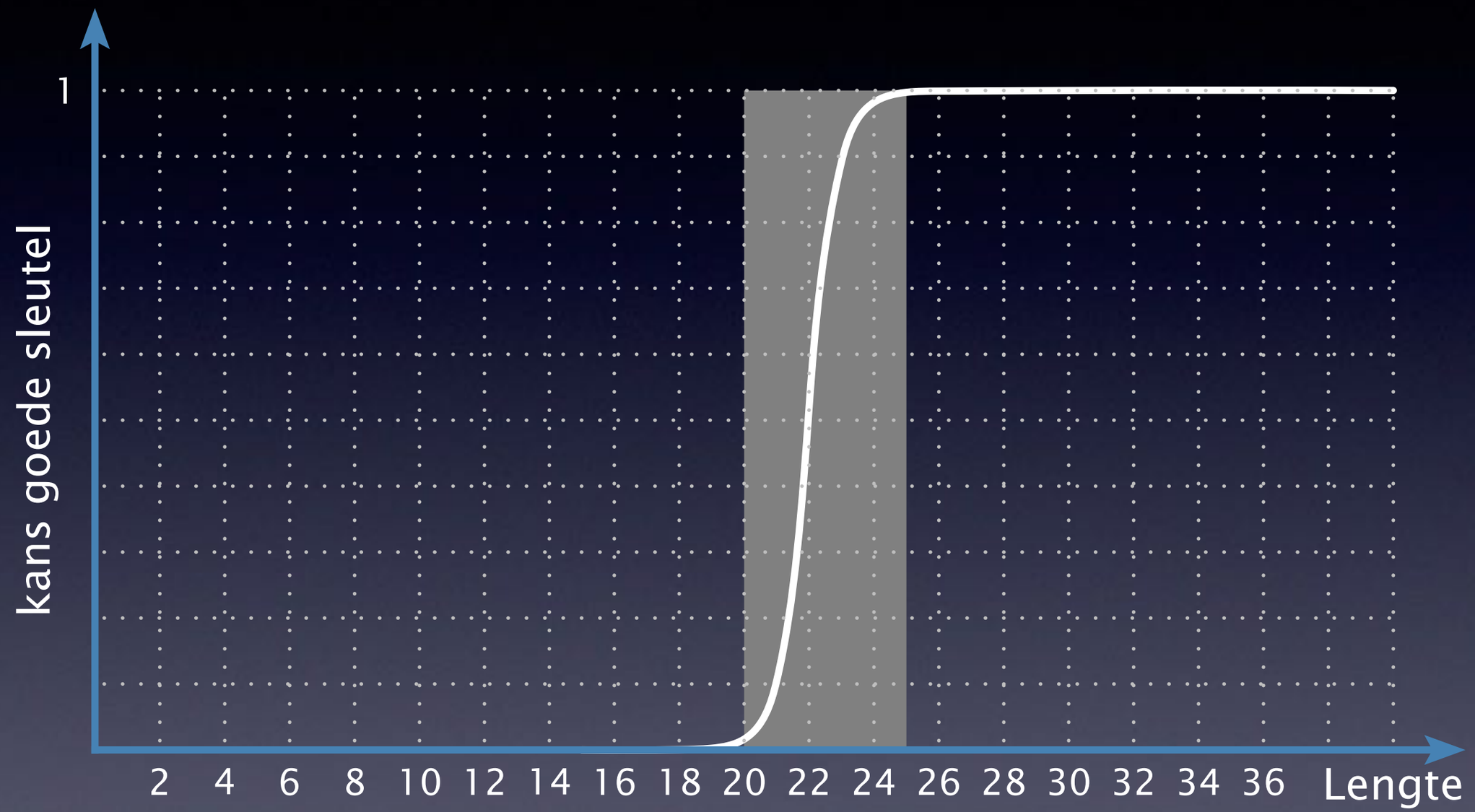
gebruik binomium van Newton $(a+b)^n$

$$P = \frac{1}{p_k |K|} (1^{|K|} - (1-p_k)^{|K|})$$

benadering omdat kans zeer klein

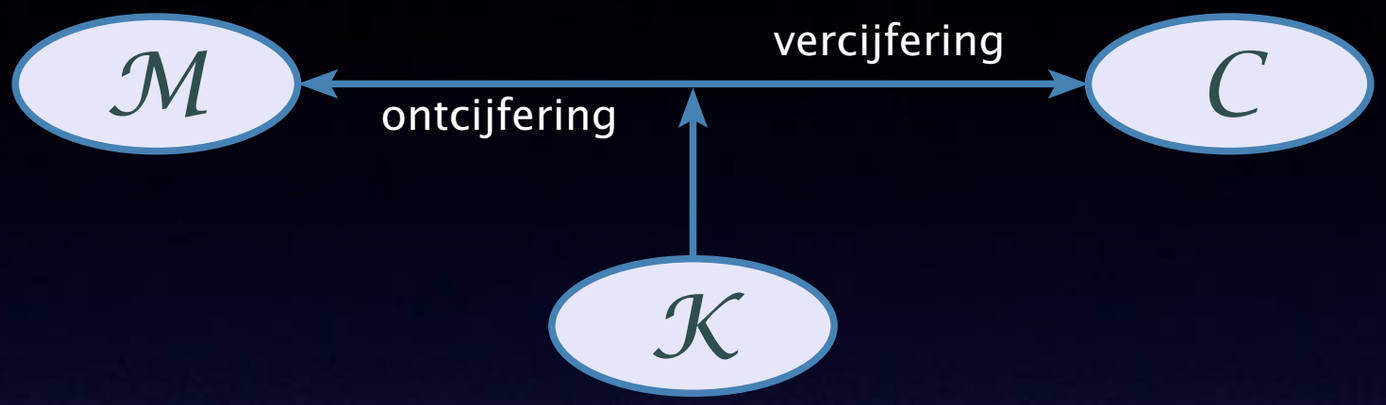
$$p_k \ll 1 \quad \rightarrow \quad P \approx \frac{1-e^{-\mu}}{\mu} \qquad \mu = p_k |K|$$

Sleutelkans



uniciteitsafstand monoalfabeet $20 \leq L \leq 25$

Veilig systeem



Regel van Bayes: $P(M|C) = \dfrac{P(M)P(C|M)}{P(C)}$ *en* $\dfrac{P(C|M)}{P(C)} = 1 \quad \rightarrow \quad P(C|M) = P(C)$