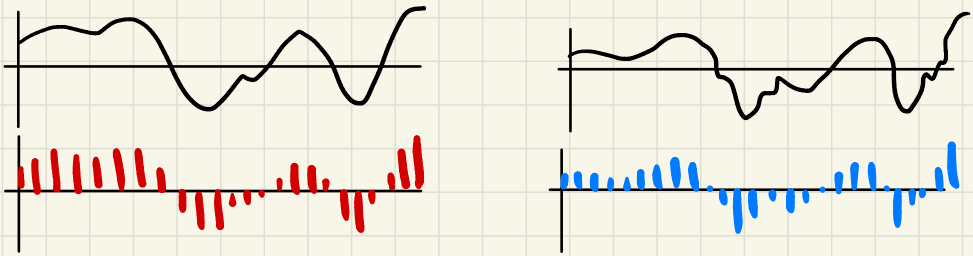
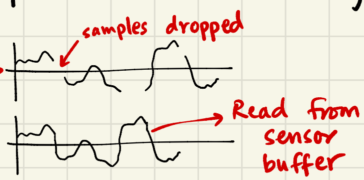
ECE/CS 434 : Dynamic Time Warping (DTW)

- Is signal s_1 similar to signal s_2 ?

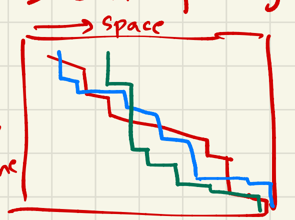


- T H E   Q U I C K   B R O W N   F O X
  T H E   Q U I C K   B R O W N   F O X

- Many real world applications go beyond delay
  - Speech — ALLEXA / AAALEXA / ALEXAAA / ALEXA
  - Typing
  - Sensor fluctuations
  - Spatial data analytics



CIA espionage

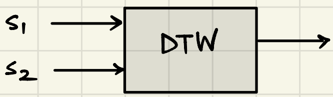


s_1 → T H E   Q U I C K   B R O W N   F O X

s_2 → T H E   Q U I C K K   B R O W N   F O X

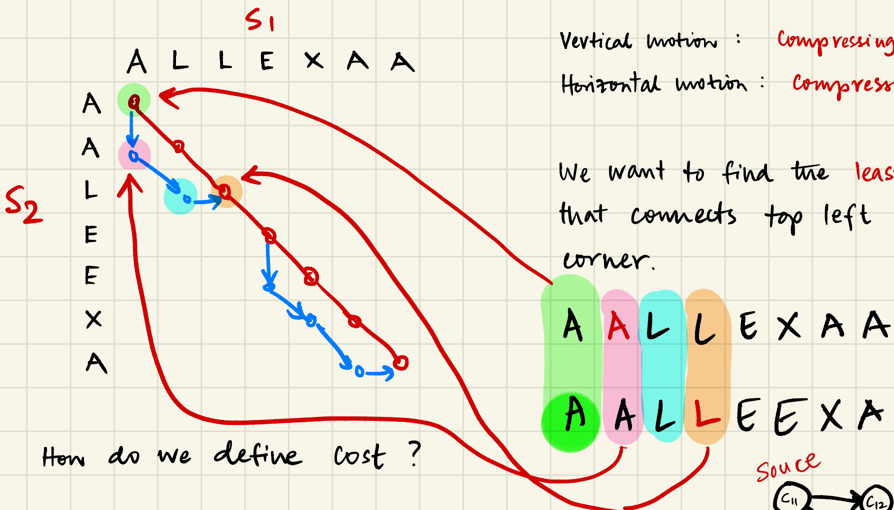
T H E E   Q U I C K   B R O W N   F O X

⊕ Dynamic Time Warping



What is the closest possible match between s_1 and s_2 by optimally stretching substrings of both signals, s_1 and s_2.

⊕ Let's model the problem for speech recognition

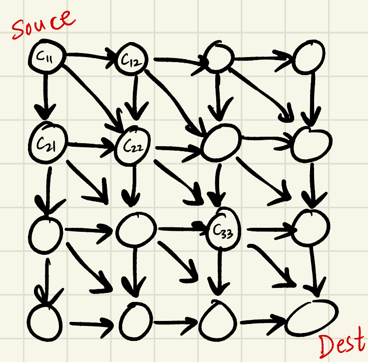


Vertical motion : Compressing vertical signal

Horizontal motion : Compressing horizontal signal

We want to find the least cost route that connects top left to bottom right corner.

⊕ How do we define cost ?



$$\text{Cost of route (R)} = \sum_{\forall (i,j) \in R} \text{node cost}(i,j)$$

Node cost $d(i,j)$ = Dissimilarity (i,j)

e.g., dissimilarity (A,A) = 0

Dissimilarity matrix :

	A	L	E	X
A	0	5	2	3
L		0	3	4
E			0	3
X				0

s_1 : A A A A L E X A

Vann s_2 : A L L L L E X A

Talal s_2 : A L E X A A A A

s_1 A A L L E E X X A A

s_2 A L E X A

⊕ DTW algorithm :

(Dynamic Program)

$$c(c_j\ r_j) = d(c_j\ r_j) + \min \{ c(c_{j-1}\ r_{j-1}),\ c(c_j, r_{j-1}),\ c(c_{j-1}, r_j) \}$$

$c(c_j\ r_j)$: Cost of going from source node to node $(c_j\ r_j)$

③

B, M

	A	L	L	E	X	A	A
A	0	5	10	12	15	15	15
A	0	5	10	12	15	15	15
L	5	0	0	3	7	12	17
E	7	3	3	0	3	5	7
E	9	6	6	0	3	5	7
X	12	10	10	3	0	3	6
A	12	15	15	5	3	0	0

Dissimilarity :

	A	L	E	X
A	0	5	2	3
L		0	3	4
E			0	3
X				0

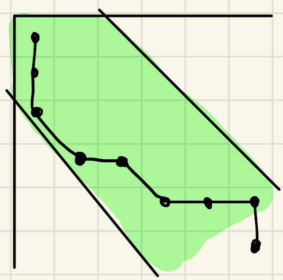
$c(1,1) = d(1,1) = 0 + \min\{c(-1,-1), c(-1,1), c(1,-1)\}$

$c(1,2) = d(1,2) = 5 + \min\{d, t, l\}$

d: diag $= 0$, t: top $= 0$, l: left $= 0$

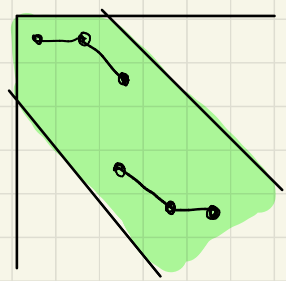
$5 + \min\{0, 0, 5\}$

③ Constraints : ①

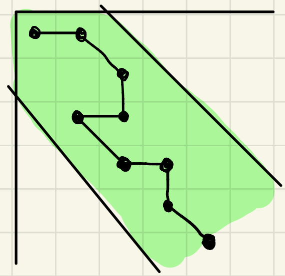


Restrict to green region, i.e., cannot compress or stretch too many in one shot.

② No discontinuity :



③ No time reversal :



DTW Takeaway

Cross Correlation → Delay
→ Noise in Signal

DTW → Insertion
→ Deletion of samples.