

THE GROWTH OF A DIRECTED WORD NETWORK

A. ROSS ECKLER

Morristown, New Jersey

A word network is a set of words and links. Two words are joined by a link if they differ by a single letter (as HORSE and HOUSE). More specifically, this is an undirected word network—one can move along a link in either direction. But what if only one direction is allowed? A simple example of a directed network is given in Chapter 4 of my *Making the Alphabet Dance*, where one changes a word to another word by removing the first letter and adding a letter at the end (as PEA to EAT, or OMIT to MITE). Here one can traverse a link in only one direction (EAT to PEA is not allowed), creating a network with many interesting properties. It is the purpose of this article to show how such a network evolves as one adds words one at a time.

Let successive words be written down on a sheet of paper and not subsequently moved, adding links as needed. The emerging network can be visualized as successive frames, one for each added word, on a computer screen. The trick is to place words so that the later patterns are clear. The clarity can be enhanced by printing words in different colors to match their function in the network at that time.

To fix ideas, look at the directed network generated by the 50 two-letter state name abbreviations used by the US Post Office. A directed link is created between two names if the second letter of the first name is the same as the first letter of the second name (as HI to IA). To show how this network develops, add names to it in the order in which states were admitted to the union:

- 1 DE appears. Call it an **isolano**, since it has no links from or to other names.
- 2-6 PA, NJ, GA, CT, MA are isolanos
- 7 MD appears with a link to DE. New terminology is needed; MD is called a **starter** (only an exit link from it), and DE is now called an **ender** (only an entrance link to it)
- 8 SC appears with a link to CT, so SC becomes a starter and CT is changed to an ender
- 9-11 NH, VA, NY are isolanos
- 12 NC appears with a link to CT, so it becomes a starter
- 13-15 RI, VT, KY are isolanos
- 16 TN is a landmark addition to the network, for it forms a cycle with NC and CT. This is the beginning of the **core**, that set of names in which any one can be reached from any other. The individual members of the core are called **insiders**. Note that CT has now played three different roles—**isolano**, **ender**, **insider**—but no further change is possible.
- 17-18 OH, LA are isolanos
- 19 IN appears with links to NY, NJ, NH and NC, and a link from RI. It is called a **joiner** since it unites RI (now a starter) with NY, NJ, NH, NC (now enders).
- 20 MS appears with a link to SC, so it becomes a starter. The SC-CT link connects a starter to an insider, so SC is christened a **preceder**.
- 21 IL has a link to LA and a link from RI, so IL becomes a joiner and LA is now an ender
- 22 AL combines with LA to form a second cycle. Note that this new section of core is physically isolated from the other part of the core. LA changes to an insider.
- 23 ME is an isolano
- 24 MO appears with a link to OH, so it becomes a starter and OH an ender
- 25 AR joins RI, IL, LA to form a cycle and a new section of core. This amalgamates with the

- LA-AL cycle to create a core section with five insiders. IN now connects two separate pieces of core, exhibiting a link from RI and a link to NC; it is called a **connector**.
- 26 MI appears with a link to IN, thus becoming a beginner; it also has a link to the insider IL
 - 27 FL appears with a link to LA, thus becoming a beginner
 - 28 TX appears with a link from CT, thus becoming an ender
 - 29 IA appears with a link from RI and a link to AL; it is an insider
 - 30 WI appears with links to IA, IN, IL, thus becoming a beginner
 - 31 CA unites the six-state core with the three-state one, exhibiting a link from NC and a link to AL. The cycle CA-AL-LA-AR-RI-IN-NC also brings IN into the core.

To summarize the state of affairs: the eleven states CA AL LA AR RI IN NC IA NC CT TN form a core in which any state can be reached from any other. Starters GA MA PA VA FL MI WI VT feed directly into the core, and starter MS accesses the core via SC. The enders NY NJ NH TX are accessible from the core. The states MO-OH and MD-DE are not connected with any of the above, nor are the isolanos ME KY.

The picture becomes clear—as one adds names, the directed network increasingly consists of an ever-enlarging core surrounded by starters and enders, and a diminishing number of starter-ender fragments and isolanos. The next increases in the core occur at frame 36 when NV sweeps VA and itself into the core via TN-NV-VA-AL, and at frame 39 when CO sweeps OR and itself into the core via NC-CO-OR-RI. When ND, ID and NM are added a new category, **follower**, is created. A follower joins an insider to an ender.

By the time all 50 states have been assimilated, the network has become too complicated to diagram. There are 30 insiders AK AL AR CA CO CT HI IA IL IN KS LA MA MI MN MO MS MT NC NH NM NV OH OK OR RI SC TN VA VT, 9 starters GA FL PA WV WI WA AK AZ UT, 7 enders NJ NY KY DE ME NE TX, 1 isolano WY and 3 followers MD ID ND.

The table below indicates the minimum number of steps needed for any insider to reach any other insider. For example, for AK to RI (or any xK to Rx) check the K row and the I column to read SCORI; the minimum number is five (AK-KS-SC-CO-OR-RI). Using this table one can more generally determine the minimum number of links joining any pair of names. Among all these minimum paths, what is the longest one? This number, known as the **span**, is six, achieved by five pairs: (IL or FL)-LA-AK-KS-SC-CT-(TN or TX) and KS-SC-CO-OR-RI-ID-DE. At frame 25, during the growth of the network, the span is seven: IL-LA-AR-RI-IN-NC-CT-TN (also starting with FL, or ending with TX starting at frame 28).

	<u>xA</u>	<u>xC</u>	<u>xH</u>	<u>xI</u>	<u>xK</u>	<u>xL</u>	<u>xM</u>	<u>xN</u>	<u>xO</u>	<u>xR</u>	<u>xS</u>	<u>xT</u>	<u>xV</u>
Ax	-	A	IA	A	SCA	A	A	CA	RIA	IA	CA	NCA	A
Cx	KSC	-	INC	NC	SC	AKSC	SC	C	KSC	INC	C	NC	TNC
Hx	RINH	OH	-	NH	SCOH	ARINH	NH	H	H	INH	COH	NH	TNH
Ix	RI	ARI	I	-	SCORI	ARI	I	HI	HI	I	CORI	NHI	ARI
Kx	K	OK	IAM	AK	-	AK	AK	CAK	K	IAK	COK	NCAK	AK
Lx	L	AL	IAL	L	SCAL	-	AL	CAL	RIAL	IAL	CAL	NCAL	AL
Mx	RINM	TNM	INM	NM	SCTNM	ARINM	-	M	RINM	INM	CTNM	NM	TNM
Nx	RIN	TN	IN	N	SCTN	ARIN	N	-	RIN	IN	CTN	N	TN
Ox	KSCO	O	INCO	NCO	SCO	AKSCO	O	CO	-	INCO	CO	NCO	TNMO
Rx	R	AR	IAR	AR	SCAR	AR	AR	CAR	R	-	CAR	NCAR	AR
Sx	KS	AKS	IAKS	AKS	S	AKS	S	MS	KS	IAKS	-	NMS	AKS
Tx	KSCT	T	INCT	NCT	SCT	AKSCT	T	CT	KSCT	INCT	CT	-	T
Vx	RINV	TNV	INV	NV	SCTNV	ARINV	NV	N	RINV	INV	CTNV	NV	-