EVERY WORD-PAIR IN THIS SET HAS ONE CRASH

A. ROSS ECKLER
Morristown, New Jersey

More than twenty years ago, in the February 1978 Word Ways, I challenged readers to find a set of eight seven-letter words in which every one of the 28 possible pairs had exactly one crash (one letter-match in the same position, as foRce and maRry, or caSinos and miSrule).

> It may be possible to locate a group of words from Webster's New International Dictionary, Second or Third Editions, but it seems far more likely that words outside these references will have to be used.

I provided a couple of near-miss examples with a set of plausible endings: -IST, -ANT, -INE, -AGE, -OGY, -ERY, -ORS, -ESS. I did not call on computers to aid in this search, a curious omission in view of the fact that Doug McIlroy had shown how they could be used to construct single 7-squares and double 6-squares in the Nov 1975 and May 1976 Word Ways. A few years later, I tried to interest various computer-literate logologists in the problem. In late 1984, Bernie Cosell, assisted by Mike Beeler, accepted the challenge, expending some 800 hours of VAX 780 CPU time (at Bolt Beranek Newman) without success. First, he perfected various programming short-cuts on the simpler problem of finding six five-letter words in which every one of the 15 pairs had a single crash, eventually discovering 104,631 solutions within the 10,000-word vocabulary. However, he discovered that the six canonical patterns into which all five-letter solutions could be classified (by suitable rearrangement of the six words) ballooned to six thousand in the seven-letter problem. Furthermore, the number of words from which eight were to be selected increased from 10,000 to 25,000. After invoking various programming tricks to save computer time, he used 60 hours of computer time to investigate three patterns--and found no solutions.

> Mike and I now both suspect that there are no solutions to be found...your casual remark about this being an instance of a "natural" for a computer because it consists of relatively simple operations vastly understates the reality. It is a huge number of operations and they are hardly "simple"...If you're hoping to do logological problems you'll have to pick your field of battle very carefully. Also, you'll have to figure programming in some pretty efficient language--forget about BASIC and FORTH.

Presented with the same problem in April 1986, Steve Root was more optimistic.

> The CLASH(C)YCLE of eight words of length seven looks as if it would yield to a couple of hours of CPU. I expect that there are at least millions of solutions. The issue is how long the computer has to throw away non-solutions..The ten words of length nine [the next more difficult problem] might take a couple days or weeks of CPU...

However, he did not write me again about it for twelve years, with an e-mail on Sep 22 1998.

> Sorry this took so long. Having a much stronger computer, a very clever program, and the Scrabble 7-letter words (start with 22281, remove the impossible Q words [to leave] 22644)...answers are slowly coming out..

And five days later, he announced he had 63 solutions, five of which consisted of words in Webster's 10th Collegiate. So much for my pessimism of 1978 and Cosell's greater pessimism of 1985!

Root found 63 solutions using 22644 words; Cosell might reasonably have expected $63(25000/22644)^8 = 134$ solutions with 25000 words. If these 134 are randomly assigned to six thousand canonical patterns, the probability that three such patterns would yield no solutions is $(.9995)^{134} = 0.935$. Cosell should not have been so quick to dismiss the possibility of a solution to the problem!

The 63 solutions are tabulated below; the ones headed by TAPIOCA, PERGOLA, BIOLOGY, SOGGILY and TANGENT are from Webster's 10th Collegiate. Note that the two SAUCILY solutions change only a single letter-crash, as do SMOKILY and MISSEAT. In the two SIMITAR solutions the swap is more complex: M,M,R in the first position convert to W,R,W; L,F,F in the third position convert to S,L,S; and M,M in the fourth position convert to P,P. All letters but Q appear in at least one solution. There is no advantage in looking for standard end-ings; every solution has a different set of three-letter endings (except for the trivial variants already noted).

Any of these solutions provides an excellent puzzle of the form "What is lexically interesting about this set of words?". Anyone who answers correctly shows an unusual sensitivity to letter-patterns in words!

TAPIOCA	PERGOLA	BIOLOGY	SOGGILY	TANGENT	SIROCCO
HETAERA	TUATARA	DEATHLY	MASONRY	PURPORT	MOMENTO
COTTONY	FULGENT	SLOSHED	BIGOTED	TEMPLAR	VINEGAR
MILIARY	PIANIST	BASTARD	REDBIRD	PODGIER	PUMICER
HIPLINE	DESTINE	SELVAGE	BARBULE	CARDIAE	MUCOSAE
TESTATE	TALCOSE	FISSILE	ROSETTE	REDBONE	PENANCE
MOSAICS	FISCALS	DALLIES	SERENES	RUNDLES	SOCAGES
CALLETS	DARNERS	FLAVORS	MIDGUTS	COMBERS	VERISTS

ZORILLO	COELIAC	COELIAC	BERSEEM	FOREARM	DARKISH
MEMENTO	METOPIC	PELITIC	TRIDUUM	CULTISM	SUCCOTH
BIBELOT	WEALTHY	HEARSAY	SLIMIER	GESTALT	SALIENT
COMBUST	GUTSILY	PIOSITY	CENTAUR	BONIEST	MISCAST
MISRULE	GOLOSHE	TOASTED	TORTILE	CINEOLE	DECIARE
BARBATE	WHEEPLE	CHOROID	CLOSURE	BERLINE	TILLITE
CASINOS	MULETAS	HILLOES	BROMALS	FUSIONS	MERLONS
ZEBRASS	CHASSIS	THEISTS	SONDERS	GILLERS	TUSKERS

FOOLISH	SALTISH	GRAVIDA	GASTREA	BEGONIA	FREESIA
TURBETH	TOWPATH	PAISANA	SENHORA	PATELLA	HOSANNA
TEENTSY	SUPPORT	TRITELY	FISHGIG	BUSTLED	FACIEND
FRAILTY	TERSEST	SUASORY	MUNTING	LANGUID	HYALOID
DUELLED	HEPTANE	PENTODE	SUBARID	CENTILE	TRAINEE
SEABIRD	BURKITE	TURDINE	FATBIRD	LISENTE	CASEOSE
SORITES	BOLSONS	SERVALS	METAGES	PIGGIES	CYCASES
DRONERS	HAWKERS	GANDERS	GIBBONS	CUTOUTS	TOELESS

GANGLIA	POGONIA	CANDELA	DIPTERA	DUSTRAG	HANDBAG
PIGNORA	SENHORA	REGMATA	BERETTA	CORDING	PEELING
FORGERY	TUGBOAT	FILMILY	PUPATED	PARTLET	SHERBET
PAUCITY	LINOCUT	SENSORY	BOLLARD	HUTMENT	CONSENT
FUNNIER	PITHEAD	SALVAGE	DELAINE	HONOREE	PAISLEY
COULOIR	SUBACID	COESITE	GUTTATE	PISMIRE	HOARILY
CIRCLES	LOBBERS	FOGDOGS	GIRLIES	CITOLAS	CHALLAS
GUGLETS	TETANUS	RIEVERS	POTEENS	DANDERS	SEIDELS

MILCHIG	GURGING	CAPELAN	POSTEEN	DRUMLIN	JALAPIN
BACHING	SANDHOG	PULSION	MILLION	CARRION	CITTERN
MYNHEER	WIDGEON	DISSEAT	HEFTILY	TORMENT	COPEPOD
BESCOUR	GOLDARN	REPAINT	PAGEBOY	SHALLOT	JEOPARD
CYCLOID	WORTHED	RUSTLER	RISIBLE	SAUTEED	TOOTSIE
DISTEND	BADLAND	CITATOR	HALITES	THYROID	PILEATE
DELLIES	SILLIES	DELETES	HALITES	DOYLIES	PETASOS
CANTHUS	BUNTERS	PATTENS	REGLETS	CRATONS	TAPPETS

OESTRIN	MISSEAT	MISSEAT	NONMEAT	RETREAT	BEIGNET
TINHORN	PARTLET	TARTLET	HELIAST	MINARET	MOONLIT
TEENIER	LUSTILY	LUSTILY	PUPILAR	PORTRAY	TRINARY
FUNFAIR	MANUARY	MANUARY	MALMIER	MUMMERY	PENALTY
EUCHRED	LOCULAR	LOCULAR	MUNDANE	ROMAINE	PROCEED
ONEFOLD	DINKIER	DINKIER	POETISE	CENTARE	TABANID
FICTILE	DORSALS	DORSALS	NEEDLES	CURRIES	MANGERS
ENSNARE	PUCKERS	TUCKERS	HAPTENS	PITMANS	BOBCATS

BULBLET	BONESET	CORSLET	PURSUIT	REDBAIT	SUNBELT
PENDANT	DISPART	MILDEST	TITLIST	BOLDEST	CHALLOT
PURSILY	PIRATED	MESSILY	SORTIED	RALPHED	COBBIER
CILIARY	COMPEND	CIPHONY	TAXPAID	BUSTARD	GENITOR
MANILLE	PENTANE	TURDINE	FIXTURE	COUTHIE	SHUNTED
BEDSORE	CASETTE	REPULSE	SUPPOSE	DESPISE	REBLEND
MIDDIES	BARTERS	RUSHEES	FATSOES	CUDDIES	ROUILLE
CARBONS	DEMASTS	TOLUOLS	POPLARS	DAUBERS	GUANINE

CATMINT	DILUENT	MIGRANT	PENNANT	SEXTANT	WHATNOT
GILBERT	SEAGIRT	CONTORT	BUSIEST	MOPIEST	SOONEST
MAMBOED	VIALLED	TINNILY	LARCENY	TIPTOED	CHOKIER
CUSTARD	DASTARD	MERCERY	POSTBOY	FORWARD	SENATOR
SULFONE	RATLINE	PYRROLE	LEGIBLE	SINCERE	PLANTED
GEMMATE	VULGATE	TOCCATE	CONCISE	TAXWISE	WEEKEND
SETTEES	SUTTEES	PECTINS	CURTALS	MERCIES	PONTINE
MISFITS	RESULTS	CYGNETS	BAGNIOS	FANIONS	CLEANSE

DOGCART	DARESAY	VALENCY	HAPPILY	LIGHTLY	PESKILY
BUNDIST	SECRECY	CURTSEY	MORTARY	NUNNERY	SORCERY
DIRTILY	CURRIED	BELLIED	TIPTOED	POTHEAD	BASTARD
NUGGETY	DISTEND	DASTARD	HUSBAND	LEEWARD	PURSUED
NARCOSE	CESURAE	DEHISCE	MISRULE	PENSILE	NOCTULE
MISDATE	LATTICE	VORLAGE	COMPONE	COGNATE	SENSATE
MONGOLS	SITUSES	BUSINGS	CURRIES	CUESTAS	NANCIES
BASTERS	LUCERNS	COHEIRS	TAMBURS	NITWITS	BUCKETS

SAUCILY	SAUCILY	SMOKILY	SMOKILY	BIGOTRY	CILIARY
BIOGENY	BIOGENY	CATTERY	MATTERY	GARGETY	MAGGOTY
BELCHED	BENCHED	OMITTED	OMITTED	BOOGIED	MISLEAD
SIALOID	SIALOID	SUSPEND	SUSPEND	LANIARD	COIGNED
TOUGHIE	TOUGHIE	OUTSOLE	OUTSOLE	GENOISE	REGINAE
CALZONE	CANZONE	DIOPTRE	DIOPTRE	COGNATE	FOSSATE
COOLIES	COOLIES	CISSIES	MISSIES	CERITES	FELLOES
TEAZELS	TEAZELS	DAIKONS	DAIKONS	LIONESS	RAISERS

SIMITAR	SIMITAR	VENULAR
MALMIER	WASPIER	BOSKIER
REFINED	RELINED	TUMBLED
COMMEND	COMPEND	DESPOND
SETLINE	SETLINE	TOLUENE
MOFETTE	ROSETTE	VAMPIRE
CATENAS	CATENAS	BALBOAS
RILLETS	WILLETS	DUNKERS

What are the chances of finding ten nine-letter words, each pair crashing only once? If 22644 seven-letter words yield 63 solutions, a simple scaling argument predicts $63/2^8 = 0.25$ solutions for half as many words, or just one solution for $11322(2^{1/4}) = 13500$ words. Steve Root has ascertained that 8538 five-letter words yield an astonishing 131,631 solutions, a typical one being HATED HORNY FITLY FAUNS WIRES WOULD. A similar scaling argument suggests that a vocabulary of 1200 words would yield, on the average, just one solution.

Solutions are ridiculously easy to find for four three-letter words. Using Kucera and Francis's *Computational Analysis of Present-Day American English* (1967) to rank-order the commonest three-letter words, one needs to examine only the 38 commonest before the solution FOR FEW HOW HER is found. By the time 127 words have been collected, sixteen solutions have appeared:

for few how her	can cut sat sun	fit far sat sir	hat how law lot
nor new how her	her hit set sir	San sex tax ten	hat him Sam sit
not new low let	bit bar sat sir	fat few law let	wet war hat her
lot led God get	bed bar had her	lie let see sit	Los law how has

Los and San appear in Los Angeles and San Francisco, respectively, in the texts Kucera and Francis sampled. On the average, about 65 words are needed for a single solution; we were a bit lucky to find it in 38.

So, extrapolating: 65 to 1200 to 13500 to ... say, 150,000 words. Don't count on finding a solution to the nine-letter problem unless you have a vocabulary several times the size of a typical unabridged dictionary!

I am much indebted to Steve Root for discovering these word groups with the aid of a computer.

SIT ON A POTATO PAN, OTIS

It's time once again for Jon Agee's delightful palindromic cartoons. Who else would think of depicting a puzzled farmer and his wife observing a gaggle of skyward-staring geese (do geese see God?), or a man with a terrified cat on his shoulder beleaguered by barking dogs, telling a neighbor what must be done (detach cat, Ed). There's more than 60 such cartoons in this $14.41 book published by Farrar Straus Giroux. If you liked *So Many Dynamos* and *Go Hang a Salami, I'm a Lasagna Hog*, this book's a winner! (ISBN 0-374-31808-5)