

what is ZoIP?

a game engine

runs over phone

virtual reality

really virtual

no consoles







no video cards





no fancy controllers





your only I/O device

is a phone





video rendered in  
your head



remember

not a game

a game engine

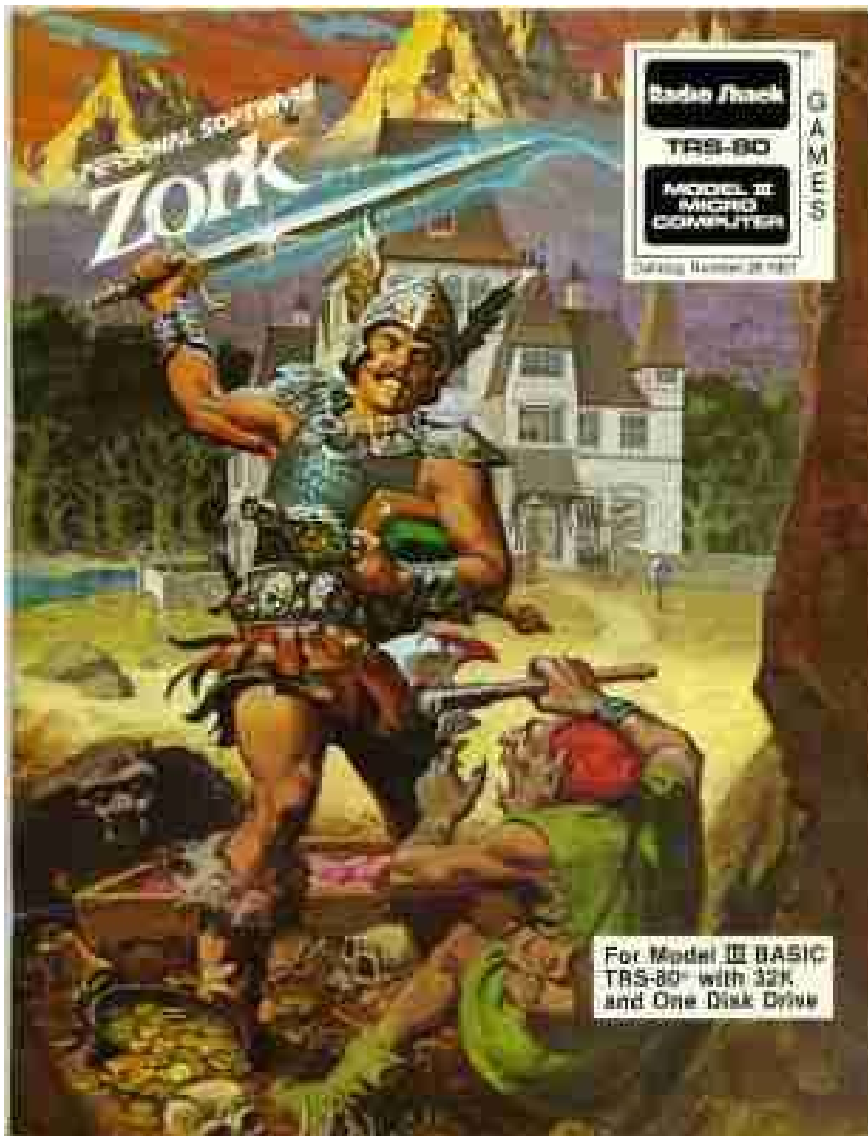
like an xbox





one game it runs

Zork



PERSONAL SOFTWARE  
**Zork**

Radio Shack  
TRS-80  
MODEL III  
MICRO  
COMPUTER

GAMES

Catalog Number 24 1807

For Model III BASIC  
TRS-80™ with 32K  
and One Disk Drive

Why Zork?

when you say...

immersive interactive  
environments for your  
telephone

???

interactive fiction



???

it's zork...

on the phone

point of reference

those new to the lore

Zork

1977

MIT



# Dynamic Modelling Group

on PDP-10





one-up

“Adventure”

**Started a company**

**Infocom**





also

makes sense

play

an epic legacy

on a legacy



so...



it's been awhile

how did it go again?

something like this...

>open trap door

The door reluctantly  
opens to reveal a  
rickety staircase  
descending into  
darkness.

>go down

You have moved into  
a dark place.

The trap door crashes  
shut, and you hear  
someone barring it.



It is pitch black. You  
are likely to be eaten  
by a grue.

>turn on lantern

The brass lantern is  
now on.

Cellar

You are in a dark and damp cellar with a narrow passageway leading north, and a crawlway to the south.

On the west is the  
bottom of a steep  
metal ramp which is  
unclimbable.

>go north

# The Troll Room



This is a small room  
with passages to the  
east & south.

Bloodstains & deep  
scratches mar the  
walls.

A nasty-looking troll,  
brandishing a bloody  
axe, blocks all  
passages out of the  
room.

The axe gets you right  
in the side. Ouch!

fast forward to the 21<sup>st</sup>  
century...

{demo ZoIP}

how did I get here?

incrementally

building simple  
applications



retrieve weather and  
read back

“Today, cloudy with a  
high of 21...”

simple DTMF

“you pressed 1, 2, 3”

found Perl z-machine

interprets Infocom  
games

like Zork

abstracted I/O



replaced print

with speak

read text

with read DTMF

version 0.1

thanksgiving



found sphinx2  
example in Perl



tested with small set

north, south, east,  
west

placed many test calls

“north...”

<tweak tweak>

“north...”

<tweak tweak>

“north...”



OMG

it worked

replaced read DTMF

with record and  
recognize

version 0.2

internals simple

1: create z-machine

2: load zork bytecode  
into z-machine



loop start:

3: create audio for  
text

4: playback audio

5: <beep>

6: begin recording

7: listen for talking

**“kill troll with sword”**

followed by silence



“ ”

■ ■ ■

8: send recording to  
sphinx2

00101011010101101  
01010110101011010

9: recognize speech

10: text is returned

<kill troll with north>

1 1: create audio file  
for what sphinx2  
thinks you said

repeat it back to you



**“You said: kill troll with  
north”**

12: type text into the  
game on your behalf

13: goto start

from there

incremental  
improvements

created large word list

no sentences

just words



didn't work so hot

modelling tools

created model with no  
grammar

without grammar

sentence structure  
unknown

“sack eat”

just as valid as

“go east”



sphinx2 has no hope

grammar needed

for long sentences

without grammar

binomial coefficients  
come into play

remember finite?

n choose k

n items



**k slots to use items**

$$n! / k! (n - k)!$$

$n!?$

**n factorial**

$$n! = (n) \times (n-1) \times (n-2) \\ \times (n-3) \times \dots (n-(n-1))$$

$$n! / k! (n - k)!$$

$$n = 200$$

$$k = 2$$



19,900

$$k = 3$$

1,313,400

$$k = 4$$

64,684,950

$$k = 5$$

2,535,650,040

grammars important



■ ■ ■

how do you make  
good grammars?

not quite sure

but like most speech  
research

it starts at CMU

<http://www.speech.cs.cmu.edu/SLM/toolkit.html>

so...

this is “cool”...



but

what does ZolP give us?

**new style IVR**

spatial cues

virtual reality

traditional IVR  
confusing

Hello, and  
welcome to  
Acme Corp



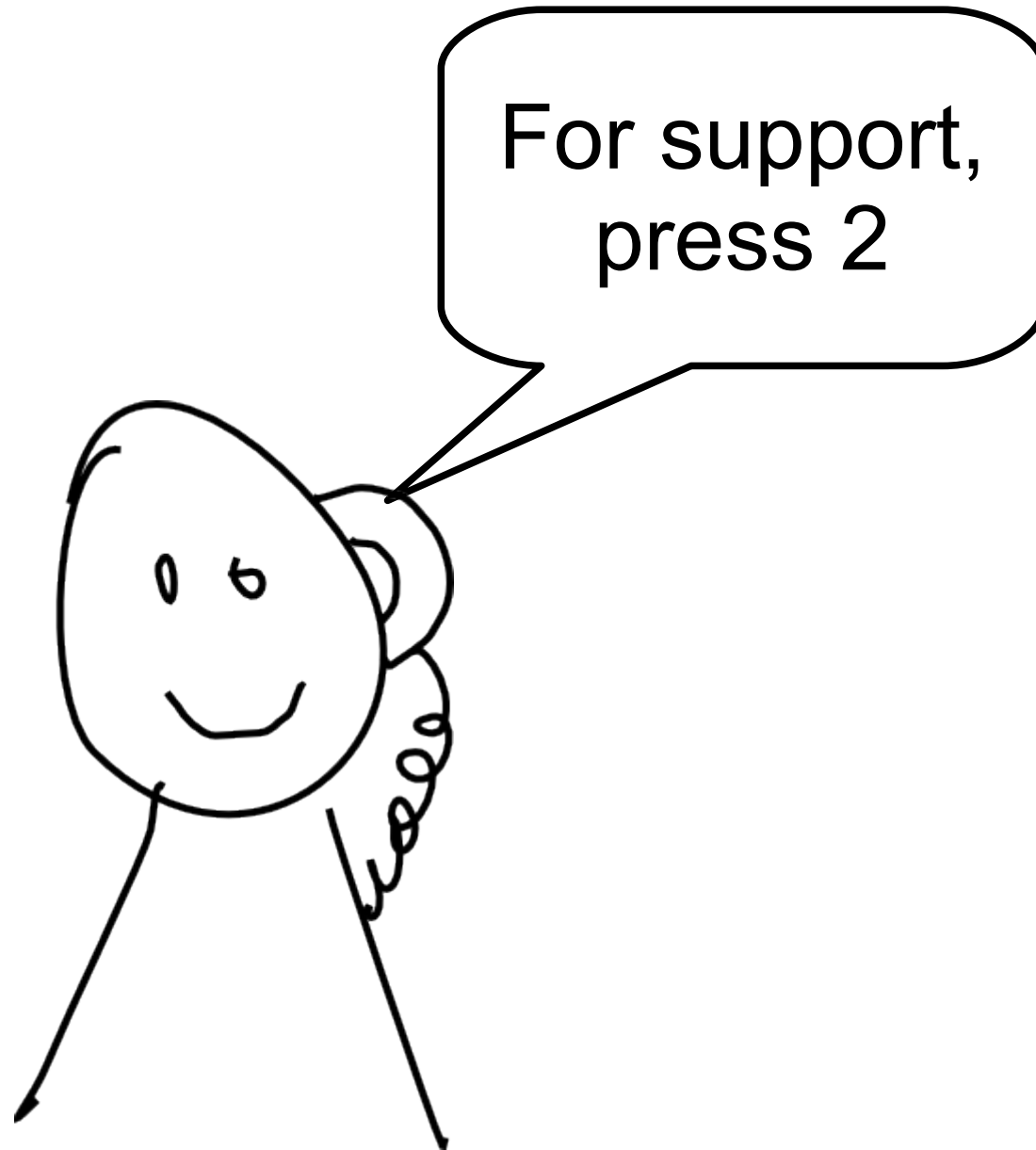
Please listen  
carefully as our  
menu options  
have changed





For sales,  
press 1





For billing,  
press 4




To check your  
account balance,  
press 6



For our corporate  
directory, please  
press #





Which was  
support  
again?

For an operator,  
please press \*

memory good for  
about 3 phrases

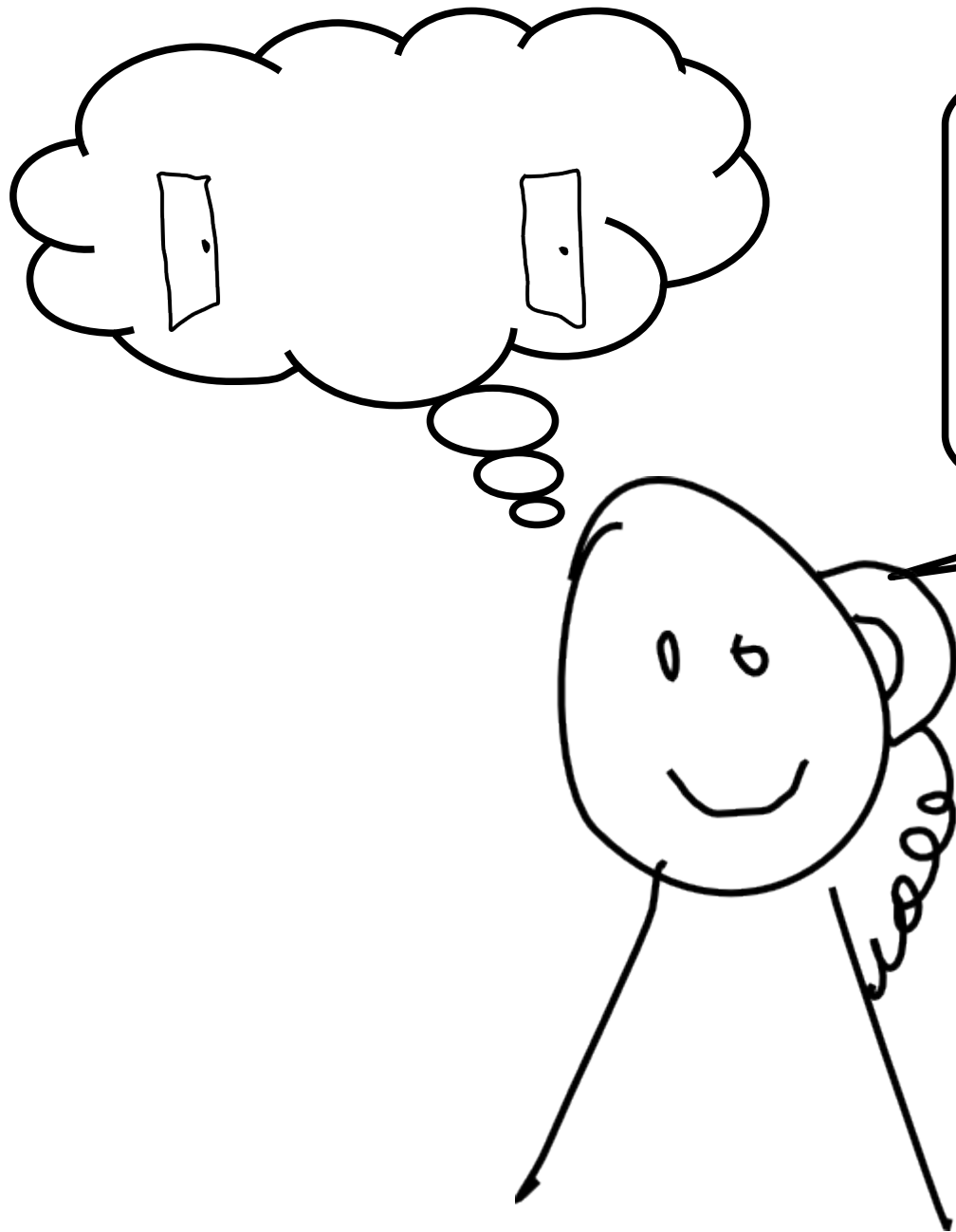
{which was support?}



spatial IVR takes  
advantage of visual  
memory

Hello, and  
welcome to  
Acme Corp





You are in the  
Lobby. To the left is  
sales, to the right is  
support



Billing is upstairs.  
There is an ATM  
here for checking  
your balance





A receptionist sits  
behind a desk,  
waiting expectantly.



visual memory good  
for 7 – 12 objects

{where is support?}

memorable?



less error prone?

allows for

richer environments

complex user  
interaction

have a sense of place

can interact with  
environment

can carry objects

can manipulate  
objects



is this better?

not sure...

what do you think?

**{fin}**

play zork now

416-548-7557

zoip.org